NTNU
Norwegian University of
Science and Technology

Modeling the Impact of Financial and Macroeconomic Variables on the Oil Price:

a VAR, Impulse Response and Markov Regime Switching Analysis

Ingvild Grøtte Bostrøm
Marina Frydenberg

Industrial Economics and Technology Management
Submission date: June 2018
Supervisor: Sjur Westgaard, IØT

Norwegian University of Science and Technology
Department of Industrial Economics and Technology Management

PROBLEM DESCRIPTION

In this study we examine how the oil price is affected by the financial factors S&P 500, VIX and Baa-Aaa credit spread, and macroeconomic variables including industrial production, interest rate and exchange rate. Their effect will be explored using well-established methods from the field of econometrics and a wide range of articles covering models used in analyzing the drivers of the oil price. First, the data will be tested to ensure that a Vector Autoregression model and a Markov regime switching model are a good fit for our time series. Then, we construct both models and analyse their findings. By applying two different frameworks, we are able to obtain a more in depth understanding of the variables' influence on the oil price.

The price of oil is often used as an indicator of global economic conditions, so understanding how the oil price behaves is a paramount from an economist's, investor's, and a policy maker's perspective. Brent crude oil has been one of the most important energy sources in the world for the last decades. According to EIA. (2018), fossil fuels will make up for 77% of the energy usage in 2040, in which oil accounts for approximately 30%. Hence, petroleum based fuels will continue to be the largest energy source for the next two decades. This means that oil is still highly relevant for the years to come. Determining the drivers of the oil price has been extensively studied in the past, but the significance of the impact of financial- and macroeconomic factors on the oil price is still being questioned.

PREFACE

This thesis concludes our Master of Science degree in Industrial Economics and Technology Management at the Norwegian University of Science and Technology (NTNU).

Since the discovery of oil in the North Continental Shelf in the 1970s, Norway has experienced an economic growth second to none, resulting in a Government Pension Fund valued at NOK 8.5 billion. The oil adventure has led petroleum to become Norway's largest industry and placed the country as the 12th largest oil exporter in the world. The importance of oil as a value creator over the last decades and for the years to come, has lead us to conduct research within this industry.

We would like to express our sincerest gratitude to our supervisor, Professor Sjur Westgaard at the Department of Industrial Economics and Technology Management, for his guidance and feedback throughout our process. Our thesis have benefited greatly from his involvement.

ABSTRACT

Oil is of great importance for the world economy, as it is the worlds largest contributor to the global energy consumption. The oil price movements are closely monitored because a sudden drop can cause ripples through the economy. A wide range of studies have been conducted on the many factors that influence the oil price, such as supply and demand, but just as important is the significance the financial- and macroeconomic factors have on the oil price.

In this thesis we address the impact of financial- and macroeconomic variables on the oil price. The time series assessed in our analysis are industrial production, interest rate, exchange rate, VIX, Baa-Aaa credit spread, S&P 500, as well as the oil price. We perform prerequisite testing for stationarity, cointegration, and structural breaks on the dataset. Based on the results from the preliminary tests, we identify a Vector Autoregressive- and Markov regime switching regression models as a good fit for this study. The Impulse response function evaluates the effects of the shocks from the variables on the oil price generated on the basis of the Vector Autoregressive model. Extending our analysis, we employ a Markov switching regression model on the time series to investigate the changes in the oil price between the high and low volatility states. Both models are adjusted to account for seasonality, stationarity and optimal lag selection.

Our results suggests that VIX, industrial production and exchange rate have the most significant impact on the oil price. VIX affects the price of oil negatively in both regimes. This finding is particularly interesting, as the relationship between VIX and oil price has not received much attention in the literature previously. We also find that the oil price responds negatively to the Baa-Aaa credit spread, the exchange rate and the S&P 500 index. Furthermore, it is evident that industrial production has a positive impact on the oil price. How the oil price responds to the interest rate is not certain, as the models show inconsistent findings.

SAMMENDRAG

Olje utgjør den største andelen av verdens energiforbruk, noe som gjør den svært viktig for verdensøkonomien. Et plutselig fall i oljeprisen kan få store konsekvenser og ringvirkninger for hele økonomien. Oljeprisens utvikling blir derfor nøye observert av analytikere verden over. Det har blitt skrevet utallige artikler om faktorer som påvirker oljeprisen, som for eksempel om tilbud og etterspørsel. Finansielle- og makroøkonomiske faktorer har derimot ikke mottatt like mye oppmerksomhet i litteraturen, og er derfor interessant å studere videre.

I denne oppgaven har vi valgt å undersøke påvirkningskraften seks finansielle- og makroøkonomiske variabler har på oljeprisen. Tidsseriene vi tar i betraktning er rentesats, dollarkurs, industriproduksjon, VIX, Baa-Aaa kredittobligasjoner, S&P 500, samt oljepris (brentolje). Vi starter med en omfattende gjennomgang av litteraturen på området for å kartlegge tidligere funn og undersøke metodene som har blitt anvendt. Deretter vil datasettet bli testet for stasjonaritet, kointegrasjon og strukturelle brudd. Basert på resultatene fra testene, ser vi at tidsseriene utfyller kravene for å utforme Vektor Autoregessiv (VAR)- og Markov regime svitsjende (MS) modeller. Modellene vil justeres for sesongeffekter, stasjonaritet og lag-lengde. Deretter vil Impulsresponsfunksjonene bli generert for å evaluere effekten sjokk tilført variablene har på oljeprisen. Videre vil vi bruke en Markov-svitsjende regresjonmodell til å evaluere oljeprisen i høy- og lav volatilitetsregimer. Funnene fra vårt studie kan bistå investorer, aksjemeglere og finansielle beslutningstakere i en bedre forståelse av oljeprisens utvikling og dens svigninger.

Resultatene våre tyder på at VIX, industriproduksjon og dollarkurs har den største innvirkningen på oljeprisen. VIX påvirker oljeprisen negativt i begge regimene. Dette funnet er spesielt interessant, da det direkte forholdet mellom VIX og oljepris ikke har vært mye studert tidligere. Vi finner også at oljeprisen responderer negativt til Baa-Aaa obligasjoner, S&P 500 og dollarkursen. Videre viser det seg at industriproduksjon har en positiv effekt på oljeprisen. Innvirkningen av rentesats er noe uklart da de to modellene gir varierende utslag på oljeprisen, til tross for dette viser begge modellene at den har en effekt.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

ACF Autocorrelation Function Plot

ADF Augmented Dickey-Fuller

AR Autoregressive

ARCH Autoregressive Conditional Heteroskedasticity

ARIMA Autoregressive Integrated Moving Average

DF Dickey-Fuller

DF-GLS Dickey-Fuller Generalized Least Squares

DW Durbin-Watson

GARCH Generalized Autoregressive Conditional Heteroscedasticity

GDP Gross Domestic Product

GNP Gross National Product

IRF Impulse Response Function

JB Jarque-Bera

KPSS Kwiatkowski-Phillips-Schmidt-Shin

LIBOR London Interbank Offered Rate

MS Markov Regime Switching

NBER National Bureau of Economic Research

OLS Ordinary Least Squares

OPEC Organization of the Petroleum Exporting Countries

Q-Q Quantile-Quantile

RW Random Walk

S.E Standard Error

USD/bbl USD/blue barrel

VAR Vector Autoregresson

VECM Vector Error Correction Model

VIX The CBOE Volatility Index

WTI West Texas Intermediate

1 INTRODUCTION

Using a multivariate Vector Autoregression model with Impulse response functions and a Markov regime switching model, this study attempts to shed light on the impact six financial- and macroeconomic variables have on the oil price. The economic factors included are industrial production, interest rate, exchange rate, VIX, Baa-Aaa credit spread, S&P 500, as well as the oil price. Understanding how they influence the oil price can assist investors, banks and portfolio managers in making important investment decisions and better predictions of the oil price.

Brent crude oil is at a current level of 76.48 USD/bbl[1], up from 46.89 USD/bbl about a year ago. According to The New York Times (2017), it will be years until the oil price recovers back to $100 a barrel, which was considered the norm until the collapse in 2014. High production quantities and low demand, were among the factors that contributed to the critical oil price drop. The continuous fluctuations have stimulated the production of a vast amount of literature on what drives the oil price.

Economists have long been intrigued by the relationship between the oil price and financial variables (see e.g., Hamilton (1983, 2003), Barsky and Kilian (2004), Akram (2009), and Bjørnland et al. (2017), to mention a few). One example is the relation between the oil price and the exchange rate. Academics are still debating the direction of causation. Fratzscher et al. (2014) and Akram (2009) propose that the causality runs negative in both direction.

There are several economic indicators that can be used to analyze the price pattern of the oil price to determine future movements. One of the indicators is the dollar because it is the benchmark for pricing oil. Measures such as gross domestic product (GDP) is also used because it reflects the economic growth in a country and is closely related to the oil price. Interest rates can give an indication of the economic conditions. High interest rates can affect the spending habits of both producers and consumers and this can have an impact on the oil price. The effect VIX, high yield bond spreads such as Baa-Aaa, and S&P 500 have on the oil price has gained more attention in the recent years, but still the literature on this subject is scarce.

[1] Oil (Brent crude) price from https://www.bloomberg.com/quote/CO1:COM accessed 06.11.2018

Moreover, the volatility should be considered as well when examining the effect of economic factors on the oil price (Oxford Energy, 2001). Oil price volatility increases the price uncertainty and the economic uncertainty overall (The World Bank, 2013). Previous studies have shown that the oil price goes through cycles of varying volatility regimes. Therefore, analyzing the financial- and macroeconomic variables and what effect they have on the oil price in different regimes can lead to more nuanced findings. Similar studies to this have been performed by Fong and See (2002), Nomikos and Pouliasis (2011) and Vo (2009). Vo (2009) finds that the oil price market has a regime like structure and emphasizes that ignoring the regime shifts will lead to a false perception that volatility is highly persistent in the oil market.

Our contribution to the literature is to supplement previous work that have been conducted on determining how interest rate, exchange rate, industrial production, VIX, Baa-Aaa, and S&P 500 affect the oil price. The purpose of this thesis is to provide market participants of information that they can use to predict the oil price more accurately. By using a different combination of variables from previous studies, this study is intended to fill the gap of how the economic factors mentioned above affect the oil price. The time series applied also contain more up to date data, which in turn makes our findings highly relevant. To make our findings more robust, we apply two different models, the Vector Autoregressive model and the Markov regime switching model. We find interesting aspects from both models that contribute to a more comprehensive study of the financial situation.

In this study we find that VIX, industrial production and exchange rate have the most significant impact on the oil price of the factors that were tested. It is worth noticing that VIX affects the price of oil negatively in both regimes. We also find that the oil price responds negatively to the Baa-Aaa credit spread, the exchange rate and S&P 500. Furthermore, industrial production has a positive impact on the oil price, and the strongest effect of all variables in the high volatility regime. The effect from the interest rate is slightly inconclusive, due to conflicting results from the models. In sum, we find that the variables have varying impact on the oil price; however, these are effects that can be explored in further research.

The remainder of the study is constructed as follows: Chapter 2 gives a discussion of the previous research conducted on the six financial- and macroeconomic drivers of the oil price. We also elaborate on past applications of the Vector Autoregressive model with Impulse response function and Markov regime switching regression (MS) used in modeling the oil market. Chapter 3 outlines the methodology for the applied tests and defines the models. Chapter 4 describes the data in detail and presents descriptive statistics and figures of the time series. Chapter 5 proceeds to present the results of the preliminary tests and then goes on to interpret the results of the models. At last, in Chapter 6, we will give our final thoughts and suggestions for further research.

2 LITERATURE REVIEW

Based on the previous literature and research, we have selected six variables that we will examine in this paper. We are interested in investigating their ability to explain fluctuations in the oil price. It is evident that financial- and macroeconomic time series have been important in understanding the current trends and behavior in the economy and will therefore be of interest when studying the movements in the oil price. In the following, we present the time series used in this study, along with the empirical literature on Vector Autoregressive (VAR) model with its corresponding impulse response function and Markov regime switching (MS) model. Lastly, we will give our contribution to the literature.

2.1 The Impact of Financial- and Macroeconomic Factors the Oil Price

In this section, we aim to give an overview of the financial- and macroeconomic variables included in this study. We briefly discuss other factors that also have shown to have an effect on the oil price, such as demand, supply and political instability, but do not take them into further consideration later in our study.

Industrial Production

One of the most important macroeconomic factors is Gross Domestic Product (GDP). In our thesis we have chosen industrial production to be a measure for this, due to the fact that it is a significant component and contributor to the GDP. Hamilton (1996) finds evidence of a positive relation between them. One of the important reasons to why the relation between industrial production and oil price should be taken into consideration, is because the industrial sector is one of the major consumers of the worlds energy. This sector has been heavily dependent on the energy supply from liquid fuels for the last decades and for the years to follow (EIA, 2017).

Exchange Rate

Another relevant macroeconomic variable is the real effective US exchange rate. In addition to being the reserve currency for most nations, the dollar is also the benchmark for pricing oil and other commodities, which is an important aspect to why we have decided to include it in our study. The bidirectional relation between currencies and commodity prices is well known in finance. As the price of the dollar strengthens, the cost of buying commodities becomes more expensive for other countries and the prices of commodities tend to fall. This finding is supported by Pindyck and Rotemberg (1988), who demonstrate that several currencies show a negative relationship with the oil price. However, in the recent years, there have been findings by Alquist and Gervais (2013) that reject any evidence of a relationship between the dollar exchange rate and oil prices. Baumeister and Kilian (2016) have also similar findings of no, or a small impact between the two variables.

Interest Rate

Even though there has been little prior evidence that interest rate movements affects the oil price significantly, it is still an important key indicator for the evaluation of macroeconomic implications on the oil price. Interest rates have shown to have an important influence on the economy, as an increase in interest rates increase the cost of borrowing, and therefore limits the spending. Interest rates are used to control inflation, and high interest rates might cause a following recession. Though there are indications that there might exist a relationship between interest rates and oil prices, the studies on the direct relationship are scarce.

S&P 500

As we try to include a wide span of variables, we incorporate an equity index in our study. There are many ways to measure the movements of the stock market, but S&P 500 is one of the most used indexes, tracked by central banks and different financial institutions all over the world. The index captures about 80% of the total market value, tracking approximately 500 of the largest stock companies in the US. The index provides us with a wide representation of the equity market and is therefore included in our study. Several studies have examined the relationship between the stock markets and the oil price, with differing findings. See studies by Bernanke (2016), Sadorsky (1999) and Hammoudeh et al. (2004).

VIX

The Chicago Board Options Exchange (CBOE) Volatility Index (VIX) measures how volatile the equity market is, derived from the implied volatility in the S&P 500 index. The VIX can be used as a proxy for market sentiment, a concept which reflects investors' attitude towards the market. When the investor sentiment is low, the market can experience raised volatility and uncertainty. As known, oil prices and other commodities can be strongly volatile at times (ECB, 2014). Ryan and Whiting (2017) also finds the VIX to be a good indicator for oil price movements.

Baa-Aaa Credit Spread

The last variable we have chosen to include in our study is Moody's Seasoned Baa-Aaa credit spread. In the past, the yield spread has been used as an indicator of economic recessions due to the expansion of the spread prior to, or during a recession. When the spread is high, investors tend to be more reserved and switch to safer assets such as AAA bonds. As a consequence, the oil price may drop. However, when the economy strengthens again, investors take on more risk, the spread narrows and the oil becomes a more attractive trade (Gertler and Lown, 1999). To our knowledge, little research has been conducted on the direct effect of the Baa-Aaa spread on the oil price. This thesis is intended to contribute more on this area.

Other factors

Other factors with an impact on the oil price are supply and demand. Historically, OPEC countries have contributed the most to the world's oil supply and therefore have a strong influence the supply, and thus the price of oil (EIA., 2018). In addition, given the significance of the region for global supply, a disruption in the area can have an impact on the fluctuations of the oil price (Kilian and Lee, 2014; Hamilton, 2003). In addition to supply, the demand can be just as influential. As the industrial production and population grows, the need for energy increases, and the oil price can experience a rise. Kilian and Hicks (2013) confirm that economic growth, as well as GDP, influences the oil price. Also, see studies by Hamilton (2008), Broadstock and Filis (2014) and Kilian and Murphy (2014). In addition, seasonality might also contribute to the changes in the oil price as the demand tends to boost in the winter season (EIA., 2018).

As discussed, oil prices are driven by a host of factors. We choose to not include supply, demand and political events in our study and focus instead on the financial- and macroeconomic variables. We refer the reader to Aastveit et al. (2015), Kaufmann (2011) and Baumeister and Kilian (2016) for more in depth study on the impact of supply and demand.

2.2 Vector Autoregressive Model and Impulse Response Function

Sims (1980) was the first to advocate Vector Autoregressive (VAR) model as an alternative to the traditional models used in econometrics. Since then, VAR has gained a lot of recognition in macroeconomic modeling. The method allows for modeling multiple variables instantaneously and interpretation of structural shocks to the VAR system by constructing Impulse response functions.

Bjørnland and Thorsrud (2015) draw attention to the difficulty to disentangle cause (i.e. the impulse) and effect (i.e. the transmission mechanism), especially regarding the US recession in the 1970's. As was first emphasized by Hamilton (1983), most of the postwar recessions in the United States have been preceded by a shock in the oil price, suggesting a negative relation between oil price shocks and the macroeconomy.

The strength of this finding, however, varies across nations (see Nkomo (2017) and Tang et al. (2010)). Since the work of Hamilton (1983), a vast number of papers have been written on the relationship between oil price shocks and the economy (see, e.g., Burbidge and Harrison (1984), Mork (1989), Bjørnland (2000a), Park and Ratti (2008), Kilian and Park (2009), Lizardo and Mollick (2010), Blanchard and Riggi (2013), and Aastveit et al. (2015)).

There is, however, limited literature on the response of the oil price to shocks in financial- and macroeconomic variables. Sadorsky (1999) tests the interaction between economic activity and oil price movements. Of particular interest to our study, is his results on the reaction of oil price to shocks to interest rate, industrial production, and stock return. Using impulse responses, he shows that an interest rate shock has a small, positive impact on the oil price, which turns down after six months. In contrast, Akram (2009) finds that the oil price falls in response to an interest rate shock.

Akram (2009) uses a Structural VAR (SVAR) model accompanied by Impulse response functions to investigate fluctuations in commodity prices. He finds that commodity prices respond to shocks in the real interest rate and the real dollar exchange rate. What can be seen from his impulse response functions, is that oil price responds positive to a shock in industrial production. However, Sadorsky (1999) finds that industrial production and stock returns have little impact on oil prices.

Fratzscher et al. (2014) examines the relationship between the oil price, the US dollar and US stock returns using Impulse response functions based on a multivariate VAR model. They find that oil prices in general react fast to a change in financial assets, such as a positive respond to a stock price shock. In response to an exchange rate shock, the oil price reacts negatively. This finding is in contrast to Akram (2009) and Sadorsky (1999), who find that a shock in the exchange rate in the short run leads to higher oil prices.

2.3 Markov Regime Switching Model

According to Granger (1996), it is evident that time series addresses regime shifts. One of the reasons for this is the indication that macroeconomic times series can experience unexpected movements such as nonlinear structural breaks, which cannot be captured by typical forecasts. According to Perron et al. (2006), this is a factor that econometric application should take this into consideration when modeling with time series. The Markov regime switching model is one of the approaches that can provide a good fit for this kind of modeling.

Application of the Markov regime switching model was early proposed by Hamilton (1989), where he examines the relationship between changes in oil prices and gross national product (GNP) growth with a utilization of the National Bureau of Economic Research recession cycles (NBER) to capture the nonlinearities in the time series. His findings suggest that oil price fluctuations play a crucial role in shocks after the postwar period and also identifies a negative relationship between the two variables. Holm-Hadulla and Hubrich (2017) also perform a Markov regime switching model to identify fluctuations in oil price shocks and their relationship with macroeconomic changes in the euro area. Ehrmann et al. (2003) incorporates the Markov regime switching model to analyze how the regime-dependent impulse responses affects the endogenous variables across different regimes. Their findings show that such an approach can reveal the associations between the regimes and periods with high or low demand on output.

Fluctuations of the oil price can also be investigated by evaluating different volatility regimes. Oil prices and commodities in general are known to have periods of high volatility. Choi and Hammoudeh (2010) demonstrate that there exist high and low volatility regimes in the oil price, other commodities and stock markets, using Markov-switching GARCH models. This finding is supported by Fong and See (2002). The results of Choi and Hammoudeh (2010) also suggest that WTI crude oil shows the strongest sensitivity to regime switching. However, shocks in the low and high volatility regimes have more persistence in the stock markets, than in oil and other commodity markets.

Research by Balcilar et al. (2015) show that there is a negative relationship between the oil price and the stock prices in the high-volatility economy and find no relationship for the low-volatility economy. There are also similar findings by Zhu et al. (2017), who identifies asymmetric effects between oil price shocks and stock returns. In other words, their findings suggest that in high-volatility regimes, structural oil shocks have significant relationship with stock returns. Another noteworthy research is conducted by Vo (2009), where he emphasizes that ignoring regime shifts in the oil market will lead to false information about the persistence of volatility regimes for the oil prices.

2.4 This Study in the Context of Existing Literature

In this thesis, we examine the effect six different financial- and macroeconomic indicators have on the oil price. The variables included in this study are the Baa-Aaa credit spread, interest rate, exchange rate, industrial production, VIX, S&P 500, as well as the oil price. This relationship will be studied using a multivariate Vector Autoregressive model with Impulse response functions and a Markov regime switching model. Therefore, a vast amount of the literature are of limited value to our study. The reasons behind this are

i) The majority of the literature focus on the impact of oil price shocks on GDP, commodity prices, interest rates, exchange rates and stock prices. The literature is scarce on how shocks in financial- and macroeconomic factors can trigger the oil price fluctuations.

ii) Our study applies a different combination of the variables in the VAR model than in previous studies, making the results from our study harder to compare to previous findings. This is due to VAR being sensitive to the number of variables included in the model, as well any omitted and included variables.

This study may be placed in the context of existing literature regarding economic indicators of the oil price. We believe this study contributes to the literature on oil price drivers, as we have an increased focus on economic indicators and include factors that have not yet received much attention in the literature, such as the VIX and the Baa-Aaa credit spread. We also apply more recent data to our study and construct two different models to provide greater insight and more robust findings.

3 METHODOLOGY

We start this section by introducing the definitions of the underlying assumptions for the VAR and MS model. Following this, we will present the models used in this study and the theory behind them. Throughout the chapter we will explain our method of choice and the reasoning behind it.

We section this part as follows. First, we define stationarity and how to account for seasonality. This lays the foundation for both the VAR- and MS model. Following this, the lag length selection criteria and cointegration test are described. Then, we will outline the VAR model and the Impulse response function. For the next part, we will present the criteria for employing a Markov regime switching model. We introduce testing for linearity, autocorrelation, heteroscedasticity, normality and breakpoints. At last, we will define the MS model with its transition- and regime probabilities.

3.1 Stationarity

Prior to estimating the Vector Autoregressive model and the Markov regime switching model, the variables must be tested for stationarity. Brooks (2008) states that all the components of the VAR system should be stationary. This also applies to the MS model. If the time series are nonstationary, this can strongly influence the behavior and properties of the time series and lead to spurious regressions.

The definition of stationarity, as described by Alexander (2008), states that a discrete time stochastic process $\{X_t\}_{t=1}^{T}$ is stationary if

1. $E(X_t)$ *is a finite constant*
2. $V(X_t)$ *is a finite constant*
3. *The joint distribution of* (X_t, X_s) *depends only on* $|t - s|$

These conditions are considered to be very formal definitions of stationarity. Therefore, we only require an independent covariance stationary, denoted as $Cov(X_t, X_s)$, which depends on $|t - s|$ and not on the total joint distribution, as described in Alexander (2008). This will be referred to as stationarity throughout the study.

We test for stationarity using four different versions of the Unit Root test. First, we perform the augmented Dickey-Fuller (ADF) test by Dickey and Fuller (1972). The ADF test is a generalization of the Dickey-Fuller test and is based on the fitting of the regression

$$\Delta X_t = \alpha + \beta X_{t-1} + \gamma_1 \Delta X_t - 1 + ... + \gamma_q \Delta X_t - q + \epsilon_t \tag{3.1}$$

The test statistic is the ratio, t, on the coefficient, β. To account for the serial correlation, the ADF test includes lags of the first differences of X_t.

Phillips and Perron (1988) propose an alternative method to the ADF test, the Philipps-Perron (PP) test, which can be expressed as

$$t_\alpha = t_\alpha \left(\frac{\gamma_0}{f_0}\right)^{1/2} - \frac{T(f_0 - \gamma_0)(se(\hat{\alpha}))}{2f_0^{1/2}s} \tag{3.2}$$

where $\hat{\alpha}$ is the estimate, t_α and t-ratio of α, $se(\hat{\alpha})$ is coefficient standard error. γ_0 is an estimation of the error variance and f_0 is an estimator of residual at frequency zero. The main difference from the ADF test is that PP is proven to be more robust when dealing with serial correlation and heteroskedasticity, without the need to specify the lag length.

As defined by Kwiatkowski et al. (1992), the Kwiatowski-Phillips-Schmidt-Shin (KPSS) test examines the reverse null hypothesis H_0: is stationary, versus H_1: has a unit root. The test statistic can be shown as the Lagrange Multiplier (LM) test against a Random Walk (RW) parameter presented below

$$KPSS = \frac{1}{T^2} \frac{\sum_{t=1}^{T} S_t^2}{\hat{\sigma_\infty^2}} \tag{3.3}$$

Here, the $S_t = \sum_{s=1}^{t} \hat{e_s}$ is the partial sum and $\hat{\sigma}_{\infty}^2$ is a HAC estimation of the $\hat{e_t}$ variance.

Lastly, we will test for unit root using the Dickey-Fuller Test with General Least Squares Detrending (DF-GLS), as proposed by Elliott et al. (1992). DF-GLS can be formulated by first detrending the series as

$$y_t^d = y_t - \hat{\beta_0} - \hat{\beta_1} t \tag{3.4}$$

where $\hat{\beta_0}$ and $\hat{\beta_1}$ can be obtained by the regression below

$$y = [y_1, (1 - \alpha L)y_2, ..., (1-)y_T] \tag{3.5}$$

$$z = [z_1, (1 - \alpha L)z_2, ..., (1 - \alpha L)z_T] \tag{3.6}$$

Here, y_t denotes the orthogonal time series and z_t denotes $[1, t]$.

The null hypothesis for this test is H_0 : there exists a unit root, and can be tested on the test statistic

$$\Delta y_t^d = \gamma * y_{t-1}^d + \sum_{j=1}^{p-1} \phi_j \Delta y_{t-j}^d + u_t \tag{3.7}$$

If the time series contain a unit root, the data is considered to be nonstationary. By taking the first difference of the data, the time series can be treated for stationarity (Bjørnland and Thorsrud, 2015). This is done by calculating the difference between two consecutive observations and then stabilizing with logarithmic transformation.

3.2 Seasonality

Time series might have a deterministic seasonality trend, which means that the series can be influenced by certain seasonality patterns (Diebold, 2015). This has previously been seen in time series such as crude oil. To correct for seasonality, monthly dummy variables are introduced and set to 1 in certain time periods, otherwise 0. Based on Hyndman et al. (2007), we choose to include the last 11 months and drop the first month. The dummy variables will be included in both the VAR and the MS model. As shown by Diebold (2015), seasonality can be included as

$$S_t = \sum_{i=1}^{s} \gamma_i D_{it} \tag{3.8}$$

where $D_{1t} = (1, 0, 0..)$ and γ_i denotes the seasonal pattern.

3.3 Cointegration

If the variables are nonstationary, one should check if they are cointegrated as well. If there is no cointegration between the variables, the VAR model can be applied. The Vector Error Correction Model (VECM) should be used instead if there exits one or more cointegrating relations between the variables. To check for cointegration, a Johansen Cointegration test as described in Johansen (1991) can be used to test the level form of the nonstationary time series.

The null hypothesis that there are at most r cointegrating vectors amounts to testing

$$H_0 : \lambda_i = 0 \text{ for } i = r + 1, ..., n \tag{3.9}$$

where only the first r eigenvalues, λ_i, are non-zero.

The Johansen test calculates the appropriate rank using the two test statistics, λ_{max} and λ_{trace}. The trace statistic, λ_{trace}, tests the null hypothesis of $H_0 : r$ cointegrating relations against the alternative hypothesis H_1 : there are more than r cointegrating relations. The trace test statistric is defined as

$$\lambda_{trace} = -T \sum_{i=r+1}^{n} \ln(1 - \hat{\lambda}_i) \tag{3.10}$$

The max statistics, λ_{max}, test the null hypothesis of H_0 : at most r cointegrating relations against the hypothesis of H_1 : there are $r + 1$ cointegrating relations. It can be defined as

$$\lambda_{max} = -T\ln(1 - \hat{\lambda}_{r+1}) \tag{3.11}$$

Bjørnland and Thorsrud (2015) explain the testing procedure as follows. The hypothesis that $r = 0$ is tested against $r > 0$ for the trace test statistic, or the alternative of $r = 1$ for the max eigenvalue statistic. The null hypothesis is rejected if the test statistic is greater than the critical value. If the hypothesis is rejected, the test are proceeded for the null of $r = 1$, and so forth. The first non-rejection is taken as the estimate of r.

3.4 Lag Length Selection Criteria

When performing VAR modeling, it is important to determine the appropriate lag length for the model. Bjørnland and Thorsrud (2015) argue that too short lag length can lead to biased estimates, whilst too large lag length may result in poor and inefficient estimates. Using a statistical criterion can help in determining the number of lags, however, in practice many choose a large lag length to begin with and check the robustness of the length thereafter. We consider three different lag selection criterion, such as the Akaike information criterion (AIC), Schwartz information criterion (SIC) and Hannan-Quinn criterion (HQC). From Liew (2004), these can be defined as

$$AIC_p = -2T[\ln(\hat{\sigma}_p] + 2p \tag{3.12}$$

$$SIC_p = \ln(\hat{\sigma^2}_p) + \frac{[p\ln(T)]}{T} \tag{3.13}$$

$$HQC_p = \ln(\hat{\sigma^2}_p) + 2T^{-1}p\ln[\ln(T)] \tag{3.14}$$

For monthly data, it is recommended to set the lag length to 12 initially, then perform different lag selection. For a dataset with more than 120 observations, Liew (2004) recommends using the Hannan-Quinn criteria for the appropriate lag length.

3.5 Vector Autoregressive Model

When the testing procedure is completed, the VAR model will be constructed. When specifying the Vector Autoregressive model, one of the main decisions is to decide which variables will be included in the system. Bjørnland and Thorsrud (2015) argue that since VAR can easily become parameterized, it is recommended to apply approximately six variables. They stress that since one has to make a decision to which variables to include, the choice should be based on economic theory. Thereafter, the lag length should be determined and non-stationary variables treated for. They remind the reader that emphasis should be put on specifying the VAR model correctly. Once the model is specified, one may estimate Impulse response functions[1] to assess the effects of a shock the system.

On a more technical note, VAR is a systems regression model and a multivariate generalization of the univariate AR(p) model. It models a linear function of each variable based on past lags of itself and the other parameters in the system. As deduced by Bjørnland and Thorsrud (2015), let a times series y at time t with ($Kx1$) vector of variables be defined as

$$y_t = (y_{1,t}, ..., y_{K,t})' \tag{3.15}$$

Then the companion form of VAR(p) model can be written as

$$y_t = \mu + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + e_t \tag{3.16}$$

where A is a (KxK) coefficient matrix, μ denotes a ($Kx1$) vector of intercept terms, and e_t is a ($Kx1$) vector of error terms, or writing out the matrices as

[1] An alternative to impulse response function is variance decompositions. It examines the VAR system of a shock to the dependent variable in proportion to shocks in the other variables (Brooks, 2008).

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{k,t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{bmatrix} + \begin{bmatrix} A^1_{1,1} & A^1_{1,2} & \dots & A^1_{1,k} \\ A^1_{2,1} & A^1_{2,2} & \dots & A^1_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ A^1_{k,1} & A^1_{k,2} & \dots & A^1_{k,k} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{k,t-1} \end{bmatrix}$$
$$+ \dots + \begin{bmatrix} A^p_{1,1} & A^p_{1,2} & \dots & A^p_{1,k} \\ A^p_{2,1} & A^p_{2,2} & \dots & A^p_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ A^p_{k,1} & A^p_{k,2} & \dots & A^p_{k,k} \end{bmatrix} \begin{bmatrix} y_{1,t-p} \\ y_{2,t-p} \\ \vdots \\ y_{k,t-p} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \\ \vdots \\ e_{k,t} \end{bmatrix} \quad (3.17)$$

As with any model, there are advantages and disadvantages. Bjørnland (2000b) mentions that VAR is easier to interpret and not so much a "black box" as other traditional large-scale macroeconometric models. On the other side, VAR has received critique for the effect the omitted variables and all types of misspecifications have on the residuals. A consequence of this is distortions in the Impulse response functions, so careful analysis should be applied to avoid overinterpretation of the results. Bjørnland also specifies that the shocks from the impulse response functions should be checked against alternative models to ensure reliability.

3.6 Impulse Response Function

VAR models are normally interpreted by the use of Impulse response functions. Impulse responses track how each variable in a VAR system react to a unit shock in a variable. For each variable, a shock can be applied to the error term, and the impulse response functions illustrate the effect of the shock on the system. This can be used to investigate how a given variable reacts to past shocks to itself, or in response to a given variable. According to Brooks (2008), the shock should eventually die out if the system is stationary.

Following his example, to illustrate how impulse responses work, let us assume the following VAR model

$$y_t = A_t y_{t-1} + u_t \tag{3.18}$$

where $A_1 = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$

For explanatory purposes, we have chosen a bivariate VAR model. The VAR can be described as

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-2} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \tag{3.19}$$

Now consider the effect of a unit shock to y_{1t} at time $t = 0$

$$y_0 = \begin{bmatrix} u_{10} \\ u_{20} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{3.20}$$

$$y_1 = A_1 y_0 = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \tag{3.21}$$

$$y_2 = A_1 y_1 = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a^2 + bc \\ ab + bd \end{bmatrix} \tag{3.22}$$

and so forth.

The ordering of the variables are important for calculating the Impulse response function. This is because the error terms are likely to be correlated, and a unit shock to the errors of one VAR equation could propagate into the system. Thus, a change in the ordering of the variables can have an effect on the results. The usual approach is to carry out orthogonalized impulse responses using Cholesky decomposition.

Bjørnland and Thorsrud (2015) state that a positive definite symmetric matrix can be written as $\sum_e = PP'$. Here, P is denoted as the Cholesky decomposition of $\sum_e$ and is represented as a lower triangular matrix. P' is the conjugate transpose. Written in terms of the VAR model in 3.18, the Cholesky decomposition can be described as

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} c_{0,11} & 0 \\ c_{0,21} & c_{0,22} \end{bmatrix} \begin{bmatrix} \nu_{1,t} \\ \nu_{2,t} \end{bmatrix} + B_1 P \nu_{t-1} + B_2 P \nu_{t-2} + ... \tag{3.23}$$

where $(\nu_{1,t})$ and $(\nu_{2,t})$ are the first and second shock, respectively.

3.7 OLS Assumptions

In this part we will introduce the tests done prior to the Markov regime switching model. The Ordinary Least Square (OLS) assumptions are evaluated for a more accurate performance of the Markov regime switching regression. Linearity, autocorrelation, homoscedasticity and normality are the assumptions needed to be considered before running the model. Failing to meet these criteria, can lead to biased and inefficient results. In addition to the evaluation of the OLS assumptions, the dataset will also be tested or structural breaks by with a Chow test.

Linearity

The first assumption implies that there should be a linear relationship present between the dependent and the independent variables. This is tested by plotting the fitted values versus the residuals. The scatter plots should have an even distribution along the the horizontal line for the linear relationship to be present.

Normality

The second assumption implies that the residuals should be close to normally distributed. It is not required for the validity of the method, but for smaller sample sizes (<200) it can be important for the calculation of p values for significance testing. To investigate this assumption, we plot the standardized residuals versus the predicted residuals. We expect to see a Quantile-Quantile (Q-Q) plot with points as close to the diagonal trend line as possible. In addition to a Q-Q plot, we graph the histograms of the time series and perform a Jarque-Bera test statistic. Jarque-Bera test (JB) can be defined as

$$JB = n\left[\frac{S^2}{6} + \frac{(K-3)^2}{24}\right] \sim \chi_2^2 \tag{3.24}$$

Here, n is the sample size, S the skewness coefficient and K the kurtosis effect. The sample is perfectly normally distributed if the JB statistic is equal to zero.

Autocorrelation

Furthermore, there should be no autocorrelation in the error terms. Autocorrealtion is often an indication of a model that is misspecified. First, we plot the autocorrelation function (ACF). The plot shows the residuals against its own lags. To investigate this assumption further, we perform Durbin-Watson (DW) test statistic, which can be described as

$$d = \frac{\sum_{i=2}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2} \tag{3.25}$$

where $e_i = y_i - \hat{y}_i$. y_i and $\hat{y}_i$ are the observed and predicted values and n is the number of observations.

The Durbin-Watson is expected to be around 2 for no autocorrelation. If the values deviate from 2, it might imply autocorrelation. If the values of the Durbin-Watson statistic are small it might indicate a positive autocorrelation, while large values might indicate negative autocorrelation.

Homoscedasticity

Lastly, we test for homoscedasticity. If the residuals are heteroscedastic, the confidence intervals are unreliable and the residuals have different variances. To test this assumption, the plot of the residuals against the fitted values is evaluated. In addition, statistic tests such as the White test and the Breusch-Pagan-Godfrey test are performed and investigated. Breusch-Pagan-Godfrey tests the serial correlation by deriving the test statistic from the errors. The formal definition of these tests can be found in Appendix C. If there is an evidence of heteroscedasticity, we correct the standard errors of the residuals by applying the HAC (Newey-West) covariance estimator, as proposed by Newey and West (1987). The elaboration of HAC (Newey-West) is described in Appendix C, section C.1.

3.8 Breakpoint test

Before completing the Markov regime switching model, the Chow test is applied to assess for initial signs of multiple regimes in our data. As denoted in Alexander (2008), this is an approach for uncovering structural breaks by using a regression model for the parameters at time t. The regression model can be described as

$$\begin{aligned} y_1 &= X_1\beta_1 + \epsilon_1, \qquad \text{for} \quad t = 1, ..., t^* \\ y_2 &= X_2\beta_2 + \epsilon_2, \qquad \text{for} \quad t = t^* + 1, ..., T \end{aligned} \tag{3.26}$$

where

$$\begin{aligned} y_1 &= (Y_1, ..., Y_{t*})', \qquad y_2 = (Y_{t*+1}, ..., Y_T)', \\ \epsilon_1 &= (\epsilon_1, ..., \epsilon_t)', \qquad \epsilon_2 = (\epsilon_{t*+1}, ..., \epsilon_T)' \end{aligned} \tag{3.27}$$

$$X_1 = \begin{pmatrix} X_{11} & X_{21} & \cdots & X_{k1} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1t*} & X_{12t*} & \cdots & X_{kt*} \end{pmatrix} \tag{3.28}$$

$$X_2 = \begin{pmatrix} X_{1,t*+1} & X_{2,t*+1} & \cdots & X_{k,t*+1} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1T} & X_{2T} & \cdots & X_{kT} \end{pmatrix} \tag{3.29}$$

Here, the X_1 ad X_2 denote the matrices. The residual sum of squares (RSS) can be calculated by estimating the model on the data up to time $t*$, then on the remaining dataset. Second, to obtain RSS_U, the two residual sums of squares should be added. Thereafter, to obtain the RSS_R, the model should be estimated with the complete dataset. There are k linear restrictions, whereby $\beta_1 = \beta_2$ is the vector of restrictions. Lastly, we use the F-test with the test statistics below to see if there exists structural breaks. This can be described as follows

$$\frac{k^{-1}(RSS_R - RSS_U)}{(T-k)^{-1}RSS_U} \sim F_{k,T-k} \tag{3.30}$$

3.9 Markov Regime Switching Model

The Markov regime switching model has shown to provide good fit and robustness when analyzing time series. It has an ability to capture significant insights in the economy over a long sample period. The Markov regime switching model, as described by Hamilton (1989), a pioneer on this field of research, allows for better modeling of the dynamics than a standard multiple linear regression. This model incorporates the assumption that the economy can follow different states and will be performed in this study to test for high- and low volatility regimes in the dependent variable, the oil price. Alexander (2008) express the model as follows

$$Y_t = \alpha_{s_t} + \beta_{s_t} X_t + \epsilon_{s_t t} \quad \text{where} \quad \epsilon_{s_t t} \sim N(0, \sigma^2_{s_t}) \tag{3.31}$$

For the two state regime, it can be written as

$$Y_t = \begin{cases} \alpha_1 + \beta_1 X_t + \epsilon_{1t}, & \epsilon_{1t} \sim N(0, \sigma_1^2), \quad \text{in state 1} \\ \alpha_2 + \beta_2 X_t + \epsilon_{2t}, & \epsilon_{2t} \sim N(0, \sigma_2^2), \quad \text{in state 2} \end{cases} \tag{3.32}$$

The regime has a latent variable, s_t, which can have two different values. They can be defined as

$$s_t = \begin{cases} 1, \text{ if state 1 governs at time t} \\ 2, \text{ if state 2 governs at time t} \end{cases} \tag{3.33}$$

The transition probabilities determine the probability of being in a state at time t, given a specific state in time $t-1$. Constant transition probabilities are assumed. The transition probability, here denoted as Π_{ij}, with the two states, i and j, can be formulated as

$$\Pi_t = \begin{pmatrix} \pi_{11} & \pi_{21} \\ \pi_{12} & \pi_{22} \end{pmatrix} = \begin{pmatrix} \pi_{11} & 1-\pi_{22} \\ 1-\pi_{11} & \pi_{22} \end{pmatrix} = \pi_{ij} \tag{3.34}$$

and the unconditional probability of regime 1 can be expressed as

$$\pi = \frac{\pi_{21}}{\pi_{12}+\pi_{21}} \tag{3.35}$$

The following vector gives the full set of parameters for the model

$$\theta = (\alpha_1, \alpha_2, \beta_1, \beta_2, \sigma_1, \sigma_2, \pi_{11}, \pi_{22})' \tag{3.36}$$

To represent the two-state regime we use a vector , ξ_t, and denote the model as

$$\xi_t = \begin{pmatrix} \xi_t^1 \\ \xi_t^2 \end{pmatrix} = \begin{cases} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{ if in state 1 at time t,} \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{ if in state 2 at time t,} \end{cases} \tag{3.37}$$

These states are considered to be unobservable, which means that we cannot be certain at which state the series are in at any time. The variable $\xi_{t|t-1}$ is introduced to express the conditional expectation of the state indicator vector, ξ_t, at time t.

The transition matrix is denoted by Π and shown as

$$\xi_{t|t-1} = E_{t-1}(\xi_t) = \Pi \xi_{t-1} \tag{3.38}$$

The probability density of the normal distribution with expectation μ and standard deviation σ can be expressed as follows

$$\phi(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} exp\left[-\frac{1}{2}\left(\frac{x-\sigma}{\sigma}\right)^2\right] \tag{3.39}$$

The equation below denotes the set of the starting value

$$\xi_{1|0} = \begin{pmatrix} \xi^1_{1|0} \\ \xi^2_{1|0} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{or} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.40}$$

The iteration proceeds as follows from $t = 1$.

1. First, it can be denoted as

$$f_t(Y_t \mid X_t;\hat{\Theta}) = \xi^1_{t-1}\varphi(Y_t;\hat{\alpha_1} + \beta_1\hat{X}_t, \hat{\sigma_1}) + \xi^2_{t-1}\varphi(Y_t;\hat{\alpha_2} + \beta_2\hat{X}_t, \hat{\sigma_2}) \tag{3.41}$$

2. Set

$$\hat{\xi}_{t|t} = \begin{pmatrix} \hat{\xi}^1_{t|t} \\ \hat{\xi}^2_{t|t} \end{pmatrix} = \begin{pmatrix} \frac{\hat{\xi}^1_{t|t-1}\varphi(Y_t;\hat{\alpha_1}+\beta_1\hat{X}_t,\hat{\sigma_1})}{f_t(Y_t|X_t;\hat{\Theta})} \\ \frac{\hat{\xi}^2_{t|t-1}\varphi(Y_t;\hat{\alpha_2}+\beta_2\hat{X}_t,\hat{\sigma_2})}{f_t(Y_t|X_t;\hat{\Theta})} \end{pmatrix} \tag{3.42}$$

3. Set $\hat{\xi}_{t+1|t} = \hat{\Pi}\hat{\xi}_{t|t}$

4. Set $t = t + 1$ before returning to step 1, while repeating the iteration until $t = T$

The iterative process provides us with:

- a set of conditional state probabilities denoted as

$$\{\hat{\xi}_{t|t}\}_{t=1}^{T} \tag{3.43}$$

- a set of conditional densities denoted as

$$\{f_t(Y_t \mid X_t; \hat{\Theta})\}_{t=1}^{T} \tag{3.44}$$

The elements of conditional state probabilities give us the conditional probability of being in state 1, or 2.

The model parameters, Θ, will be estimated by maximizing the log likelihood function

$$\ln L(\Theta) = \sum_{t=1}^{T} \ln f_t(Y_t \mid X_t; \Theta) \tag{3.45}$$

It is performed a sub-iteration for each iterative step. By using the estimated model parameters, this maximization gives us the conditional densities and set of conditional state probabilities. We will examine the coefficients obtained from the model in the results chapter.

4 DATA ANALYSIS

In this section we will perform data analysis of the time series. To begin with, the descriptive statistics, correlation matrices and time series plots are presented for each variable. The time series are also tested for stationarity, a prerequisite for the Vector Autoregressive model and the Impulse response function. Next, we provide descriptive figures to evaluate the OLS assumptions, which are later taken into consideration prior to performing the Markov regime switching model.

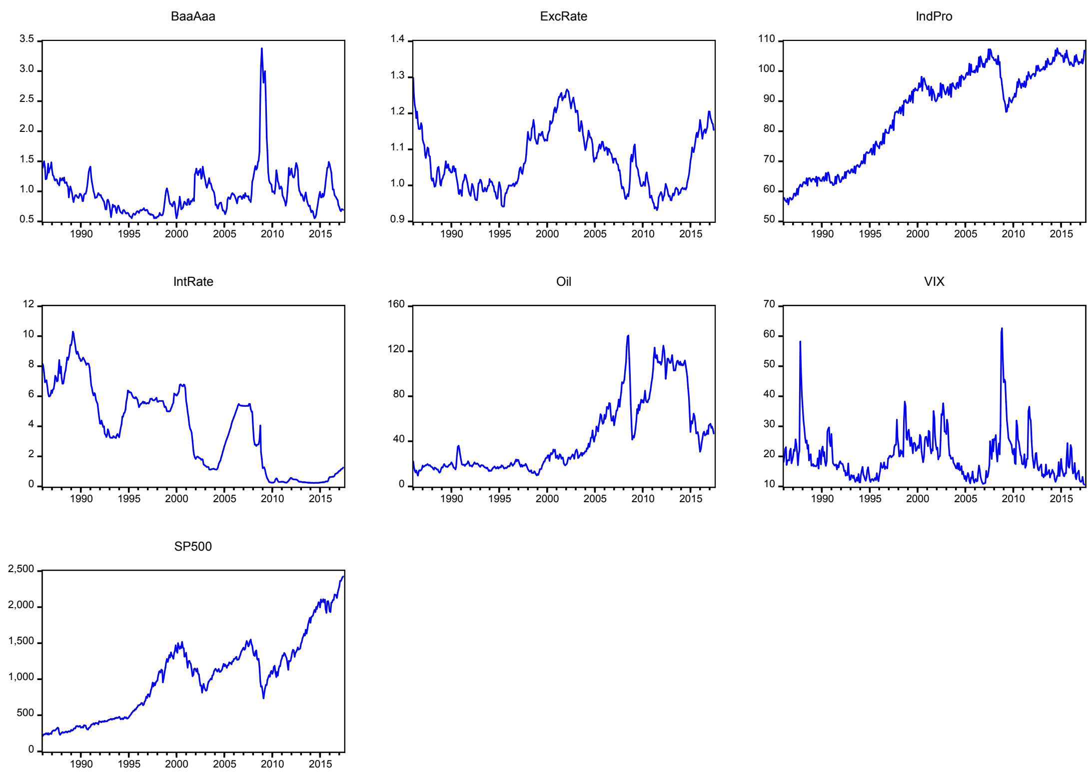
4.1 Descriptive Statistics

The financial databases used in this paper are the Federal Reserve Economic Data (FRED) and Yahoo Finance. Both offer detailed economic data, updated on a regular basis. We have obtained monthly data from 1986M01 to 2017M06 of seven variables, which will be presented shortly. One of the time series, VIX, is gathered from Yahoo Finance, whilst the other time series are obtained from FRED.

The time horizon is selected to comprise influential events for the oil price in the recent years, with the interest of capturing any indications of structural breaks. The time horizon chosen comprises important events such as the Iraq war, 9/11, the world financial crisis in 2008, and the great oil crash in 2014.

The plots in figure 4.1 display the variables over the time period 1986M01-2017M06. BaaAaa is the equity spread of Moody's Seasoned Baa Corporate Bond Yield and Moody's Seasoned Aaa Corporate Bond Yield. IndPro is the Industrial Production: Total index. IntRate is the 3-Month London Interbank Offered Rate (LIBOR), based on U.S. Dollar. Oil is the global price of Brent Crude. SP500 is Standard & Poor's 500 index. ExcRate is the Real Effective Exchange Rates Based on Manufacturing Consumer Price Index for the United States. VIX is composed of the Chicago Board Options Exchange (CBOE) Volatility Index starting at 1990M1, and the CBOE S&P 100 Volatility Index (VXO) from 1986M01 to 1989M12 due to VIX not having data prior to 1990. Figure A.1 displays the normal distributions for each variable and A.2 displays the scatter plots for each of the six variables against the oil price and can be seen in Appendix A.

Figure 4.1: Time Series Plots



Time series of all seven variables in levels. From the left starting with BaaAaa (Baa-Aaa), ExcRate (Exchange rate), IndPro (Industrial Production), IntRate (Interest rate), Oil (Oil price), VIX (VIX) and SP500 (S&P500). The sample period is 1986M01-2017M06.

Table 4.1 reports the descriptive statistics of the time series. Key takeaways from the table is the high volatility of the oil price. This can have consequences for accurate interpretation of market movements. The series have a mean around zero and medians that deviate from the mean. This can be an indication of non-normally distributed data, a typical characteristic for commodities and other financial assets with an indication of fat-tail distribution. This is supported by the kurtosis and Jarque-Bera statistics. All of the variables have kurtosis larger than 3, which implies that there might be more extreme values than in a normal distribution.

Next, we examine the correlation between the oil price and the economic variables for the data in level form and first differences, which are reported in table 4.2 and 4.3, respectively. We observe that oil price has positive correlation to industrial production and S&P 500. The oil price has a negative correlation to interest rate and exchange rate. The observed correlation between oil price and VIX is negative. The correlation to Baa-Aaa are positive in levels and negative in first differences.

Table 4.1: Descriptive Statistics

	BaaAaa	ExcRate	IndPro	IntRate	Oil	SP500	VIX
Mean	-0.001858	-0.000314	0.001638	-0.004950	0.001986	0.006465	-0.001445
Median	-0.010929	-0.000400	0.000744	-0.000235	0.005902	0.011035	-0.009146
Max	0.448694	0.059001	0.054724	0.386250	0.466204	0.123780	0.980028
Min	-0.287682	-0.042467	-0.050752	-0.577070	-0.313455	-0.245428	-0.372980
Std.Dev	0.079776	0.012791	0.019395	0.083990	0.092711	0.043829	0.153408
Skewness	0.854594	0.083678	0.016664	-1.310157	-0.071010	-1.082592	1.444171
Kurtosis	7.585789	4.085933	3.377835	13.59450	5.290197	6.717792	9.090048
Jarque-Bera	376.2269	18.96403	2.259954	1871.012	82.70709	290.7612	713.6485
Probability	0.000000	0.000076	0.323041	0.000000	0.000000	0.000000	0.000000
Observations	377	377	377	377	377	377	377

Descriptive statistics of the log returns of the all the seven time series, BaaAaa (Baa-Aaa), ExcRate (Exchange rate), IndPro (Industrial Production), IntRate (Interest rate), Oil (Oil price), VIX (VIX) and SP500 (S&P500). The sample period is 1986M01-2017M06.

Table 4.2: Correlation Matrix in Levels

Correlation	BaaAaa	IndPro	IntRate	Oil	SP500	ExcRate	VIX
BaaAaa	1.000000						
IndPro	0.015302	1.000000					
IntRate	-0.178154	-0.689536	1.000000				
Oil	0.164335	0.670015	-0.643554	1.000000			
SP500	-0.074619	0.913349	-0.663921	0.596245	1.000000		
ExcRate	0.114177	0.188153	0.025605	-0.379504	0.215949	1.000000	
VIX	0.619599	-0.038187	0.066995	-0.078959	-0.167177	0.219374	1.000000

Correlation matrix for all seven time series stated in levels, BaaAaa (Baa-Aaa), ExcRate (Exchange rate), IndPro (Industrial Production), IntRate (Interest rate), Oil (Oil price), VIX (VIX) and SP500 (S&P500). The sample period is 1986M01-2017M06.

Table 4.3: Correlation Matrix in First Differences

Correlation	BaaAaa	IndPro	IntRate	Oil	SP500	ExcRate	VIX
BaaAaa	1.000000						
IndPro	-0.085691	1.000000					
IntRate	0.209265	0.093373	1.000000				
Oil	-0.212448	0.063752	-2.13E-05	1.000000			
SP500	-0.156076	-0.089004	-0.100376	0.022529	1.000000		
ExcRate	0.178164	0.115059	0.238122	-0.219741	-0.242145	1.000000	
VIX	0.199640	0.148365	0.205619	-0.087256	-0.555666	0.169240	1.000000

Correlation matrix for all seven time series stated in first differences, BaaAaa (Baa-Aaa), ExcRate (Exchange rate), IndPro (Industrial Production), IntRate (Interest rate), Oil (Oil price), VIX (VIX) and SP500 (S&P500). The sample period is 1986M01-2017M06.

4.2 Stationarity Transformation

An important criteria for the VAR model is that the variables have to be stationary (Bjørnland and Thorsrud, 2015). Four different Unit Root tests are applied to the time series to test for this assumption, and a complete list of the test results are presented in table A.1 and A.2 in Appendix A. The Unit Root tests are first reported for the time series in levels and second for logarithmic first differences, both for 1%, 5% and 10% test statistic.

For ADF and DF-GLS, the lag length was set according to the Schwarts Info Criterion with maximum lags set to 16. As observed in the table A.1, the augmented Dickey-Fuller test fails to reject the null hypothesis of a Unit Root for the variables, except for Baa-Aaa and VIX. Thus, indicating that the other five variables are nonstationary. This is confirmed by the other Unit Root tests. All of the variables are therefore transformed into logarithmic differences, which should correct for nonstationarity (Bjørnland and Thorsrud, 2015). Further, the variables are tested again for Unit Roots, but this time in first differences. The results are recorded in table A.2. The tests show that all of the variables are now stationary.

4.3 OLS Assumptions

We examine the set of OLS assumptions before performing the Markov regime switching model. The OLS assumptions need to be fulfilled in order for the regression to perform correctly. The most important assumptions are linearity, no serial correlation, normally distributed- and homoscedastic errors. The set of assumptions are evaluated against the residuals from a least square regression.

To start with, the linear relationship between the dependent variable, oil price and the independent variables, Baa-Aaa, industrial production, exchange rate, interest rate, VIX and S%P 500, are evaluated by studying the plot of the residuals versus fitted values, presented in figure 4.2(b). Most of the observations are close to evenly distributed, with a few deviating observations. These are considered as white noise, which can be due to unexpected events or factors affecting the oil price. Thus, the relationship seems to be linear.

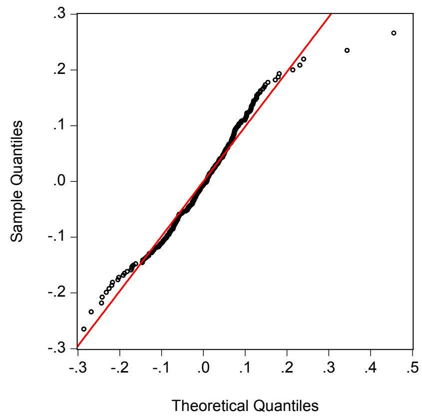
Next, the second assumption of no serial correlation is investigated by performing the Durbin-Watson test statistic, as seen in table A.5. The test statistic shows approximately 1.66, which indicates that there is a slight presence of autocorrelation in the residuals. Further, the ACF plot in figure 4.2(c) is investigated. The vertical lines shows the value of the autocorrelation for the specific lag value. As it can be seen, the serial correlation occurs at several lags at a 5% significance level. Specifically, it can be observed positive autocorrelation at lag 1, negative between lag 3-5, 13-17, and at lag 20. However, it stays within the significance range of the red, dashed line.

To evaluate the normality assumption, both the Q-Q plot and the Jarque-Bera test statistic are performed. From the figure 4.2(a), the residuals seem to lay along the fitted line, except from a few outliers, which can be an indication of fat tails. The Jarque-Bera test statistic of approximately 94.45 also supports this finding, see figure 4.2(d) with the corresponding histogram. The value of kurtosis, 5.44, seems to be slightly high, even though the skewness of 0.069 is close to zero. Since normality is not strongly required, we continue with the analysis.

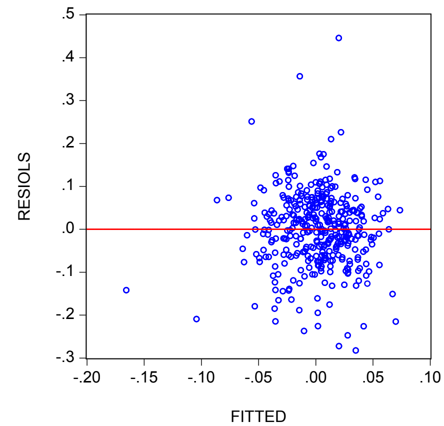
Lastly, several tests are performed to investigate whether the errors are homoscedastic. The results of the White test and Breusch-Pagan-Godfrey test can be found in Appendix A, table A.3 and A.4, respectively. The tests reject the null hypothesis of homoscedasticity, as it can be seen from both F-statistics. If the errors are not homoscedastic, then this is an evidence of unreliable confidence intervals. To correct the standard errors, we apply the HAC (Newey-White) method. The original residuals and corresponding standard error are presented in table A.5 and the corrected standard errors are presented in the table A.6 in Appendix A.

Figure 4.2: OLS Assumptions: Descriptive Figures

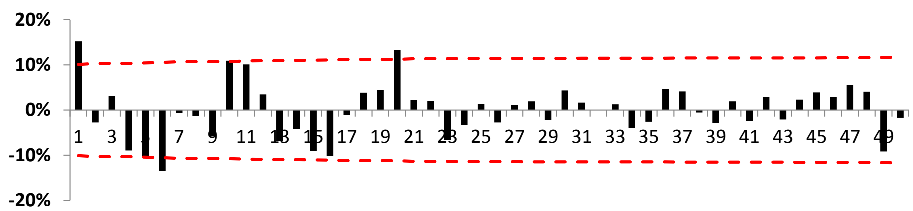
(a) Q-Q Plot



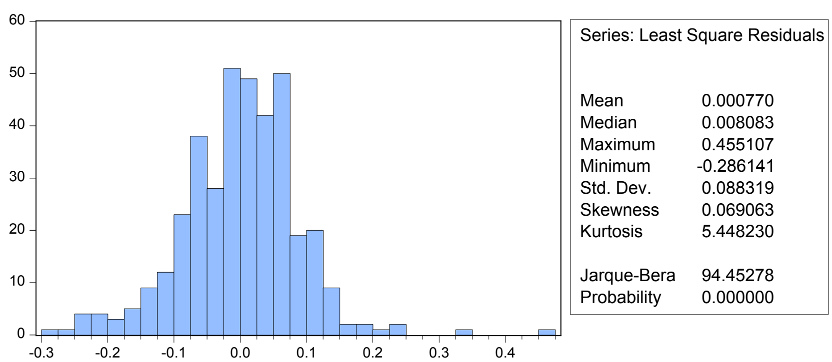
(b) Fitted vs Residuals Plot



(c) ACF Plot



(d) Histogram



Descriptive figures for the residuals of the least-squares equation for oil as the dependent variable and BaaAaa, industrial production, interest rate, exchange rate, VIX and S&P500 as the independent variables, in first differences. Sample period: 1986M01-2017M06. (a) Q-Q plot. The linearity of the points suggests that the data are normally distributed. Outliers are visible in the upper right and lower left corners. (b) Fitted vs residuls plot. Close to evenly distributed around the horizontal line. White noise in upper right corner and lower left.(c) ACF plot includes 50 lags at 5% significance. (d) Histogram of Least Square residuals.

5 RESULTS

The following section discusses the results from the modeling and is structured as follows. At first, the preliminary tests of the VAR model are conducted. The time series are tested for cointegration, and the appropriate lag length is determined. Based on the results, a VAR model is constructed and impulse response functions of the oil price to shocks to the financial- and macroeconomic variables are generated and then evaluated. For the next section, the time series are tested for structural breaks. This leads to the second model in this thesis, the Markov regime switching model. At last, the results from the regime switching model are reviewed.

5.1 Cointegration Test

The number of cointegrating relationships is determined by performing the Johansen cointegration test. The nonstationary variables in level form will be tested to check whether there exists a long-run causal relationship between the variables. The test is performed to evaluate whether to use a VAR or a VECM model. The latter is designed for nonstationary series, which are cointegrated (Brooks, 2008). If there are no cointegrating relationships, a VAR model will be a better fit for the time series.

The Johansen test and its results are reported in table 5.1. The nonstationary variables tested are interest rate, exchange rate, industrial production, oil price and S&P 500. The table shows the λ_{trace} and λ_{max} statistics, with their corresponding 5% critical values. The null hypothesis of no cointegrating relation is not rejected against the alternative hypothesis of $r > 0$ at the 5% critical level for the trace test. The max eigenvalue rejects the null against the alternative of $r = 1$, and has its first non reject for the null hypothesis of $r \leqslant 1$ versus the alternative hypothesis of $r = 2$. However, since the first non reject was for the trace test for the null of no cointegrating relations, we proceed with the VAR model in this study.

Table 5.1: Johansen Cointegration Test

H_0	λ_{trace}	5% critical value	λ_{max}	5% critical value
$H_0 : r = 0$	61.31046	69.81889	35.33508	33.87687
$H_0 : r = 1$	25.97538	47.85613	13.55261	27.58434
$H_0 : r = 2$	12.42277	29.79707	8.306213	21.13162
$H_0 : r = 3$	4.116559	15.49471	4.116444	14.26460
$H_0 : r = 4$	0.000115	3.841466	0.000115	3.841466

Johansen Cointegration Test for the nonstationary variables in levels with the report of trace (λ_{trace}) and maximum eigenvalue (λ_{max}) test statistics with their corresponding 5% critical value, sample period is 1986M01-2016M06.

5.2 Lag Length Selection Criteria

We employ a multivariate information criterion presented in table 5.2. Initially, the lag length is set to 12. The Hannan-Quinn Criterion (HQC) is favored when the number of observations is 120 or more (Liew, 2004). From the table, HQC suggests using a lag length of 1 for the VAR model. This is also confirmed by the Akaike information criterion (AIC) and Schwarz criterion (SIC).

Table 5.2: Lag Length Selection Criteria

Lag	AIC	SIC	HQC
0	-23.93685	-23.03934	-23.58016
1	-24.63130*	-23.21024*	-24.06655*
2	-24.58710	-22.64249	-23.81429
3	-24.51903	-22.05087	-23.53815
4	-24.47589	-21.48419	-23.28695
5	-24.34895	-20.83370	-22.95195
6	-24.25099	-20.21219	-22.64592
7	-24.12007	-19.55772	-22.30693
8	-24.02094	-18.93505	-21.99974
9	-23.92998	-18.32054	-21.70072
10	-23.87870	-17.74571	-21.44137
11	-23.88912	-17.23258	-21.24373
12	-24.02967	-16.84958	-21.17621

Lag length selection criteria with the report of AIC, SIC and HQC information criteria. Performed on 12 initial lags, for the seven variables (in first differences). Sample period of 1986M01-2017M06.

5.3 Vector Autoregressive Model

The multivariate VAR model is estimated with monthly data from 1986:01 to 2017:06. The time series are in first difference, the lag length is set to 1, and seasonality effects are taken into account. The dependent variable is the oil price, the independent variables are the Baa-Aaa credit spread, interest rate, exchange rate, industrial production, VIX and S&P 500. In what follows, the results of the VAR model reported in table 5.3 will be examined. The columns in the table correspond to equations in the VAR system, and the rows corresponds to the regressors in the equations. The VAR equations and corresponding coefficients are reported in table B.1 and B.2, Appendix B.

The oil price receives positive and statistically significant influence from the lagged value of Oil(-1). Furthermore, the lagged value of the ExcRate(-1) (exchange rate) has a negative and statistically significant impact on the oil price, but this relationship is not bidirectional. Lagged values of IndPro(-1) (industrial production) , IntRate(-1) (interest rate), SP500(-1) (S&P 500), VIX(-1) (VIX) and BaaAaa(-1) (Baa-Aaa spread) are not statistically significant.

Table 5.3: Vector Autoregression Estimates

	Oil	IndPro	IntRate	ExcRate	SP500	VIX	BaaAaa
Oil(-1)	0.181182	0.022723	0.060949	0.002487	-0.008035	-0.000284	-0.084005
	[3.43597]	[4.40649]	[1.43767]	[0.35602]	[-0.30584]	[-0.00320]	[-1.98295]
IndPro(-1)	0.585311	-0.048353	0.056580	-0.163106	0.477122	-0.349712	-0.696871
	[1.10619]	[-0.93447]	[0.13300]	[-2.32644]	[1.80975]	[-0.39275]	[-1.63934]
IntRate(-1)	0.027563	0.015534	0.562317	0.009198	-0.024368	0.280249	0.134353
	[0.46386]	[2.67334]	[11.7708]	[1.16831]	[-0.82305]	[2.80269]	[2.81440]
ExcRate(-1)	-1.033653	0.034903	0.205429	0.377029	-0.176910	-0.540623	0.066776
	[-2.60439]	[0.89928]	[0.64380]	[7.16940]	[-0.89460]	[-0.80945]	[0.20942]
SP500(-1)	-0.120750	0.029366	0.054167	0.010414	-0.004785	-0.570004	-0.297178
	[-0.94049]	[2.33889]	[0.52477]	[0.61216]	[-0.07480]	[-2.63822]	[-2.88109]
VIX(-1)	-0.039127	0.007953	-0.042841	0.001365	-0.022864	-0.131639	0.045679
	[-1.02466]	[2.12975]	[-1.39546]	[0.26975]	[-1.20168]	[-2.04855]	[1.48898]
BaaAaa(-1)	-0.089906	-0.010495	-0.109614	-0.000105	0.013234	-0.138334	0.239095
	[-1.43568]	[-1.71372]	[-2.17719]	[-0.01269]	[0.42414]	[-1.31270]	[4.75240]

The table reports the estimates from the VAR model based on a sample period of 1986M01-2017M06. The columns represent the equations, and the rows reports the lagged parameters. The variables are denoted as BaaAaa (Baa-Aaa), IndPro (industrial production), IntRate (interest rate), SP500 (S&P 500), ExcRate (exchange rate), VIX(VIX) and Oil (oil price). The t-statistic is denoted as [...]

5.4 Impulse Response Function

Based on the VAR model, the impulse responses are traced out. In the following section, an analysis of the Impulse responses of oil given a shock to industrial production, interest rate, exchange rate, VIX, Baa-Aaa credit spread, S&P 500, and oil price is given.

As this approach is not indifferent in the ordering of the variables, the Impulse response function in this thesis is calculated using the ordering *industrial production, interest rate, exchange rate, VIX, Baa-Aaa credit spread, S&P 500, oil price.* Given the focus of our study on how the economic factors affect the oil price, we have chosen the place the economic factors first and the oil price last. This is consistent with the work of Akram (2009), where he includes commodity prices last. The ordering of industrial production, interest rate, and exchange rate is used in both Akram (2009) and Eichenbaum and Evans (1995).

The impulse responses can be viewed in figure 5.1, and the numerical values of the oil after shock the economic variables are reported in table B.3, in Appendix B. For the curious reader, the impulse response functions of all the financial- and macroeconomic factors are reported in Appendix B, figures B.1 through B.6.

The oil price has an initial positive reaction to a shock in industrial production, similar to the findings of Sadorsky (1999) and Akram (2009). The effect lasts about six months. The strongest positive influence takes place withing two months of the shock. A possible interpretation of this finding is that when industrial production increases, the demand of oil increases and consequently, the oil price rises.

Taking a closer look at the impulse response of the oil price to an interest rate shock, it is evident that the shock has a negative impact on the oil price. The shock appears to have faded out of the system after seven months. This result is fairly consistent with the pattern observed by Akram (2009). A common belief is that an increase in interest rates raise costs for consumers and manufacturers, which in turn reduces the demand of oil. Thus, the oil price decreases.

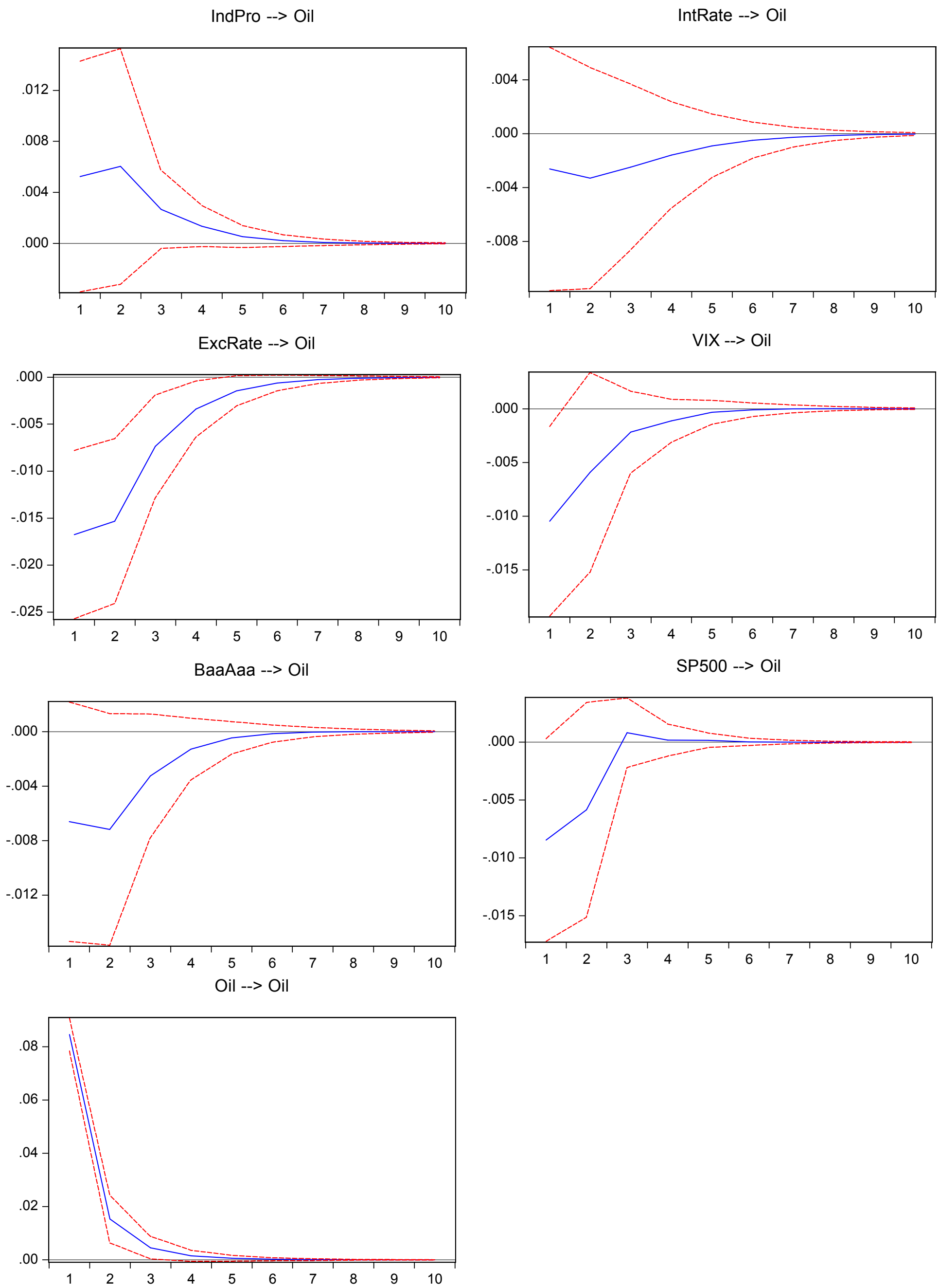
We observe that a shock to the exchange rate has a negative and statistically significant impact on the oil price. This effect dies out after about seven months. A possible explanation of the relation is that commodities become costlier in other currencies when the dollar strengthens. Thus, the demand for oil decreases, and the oil price drops. Our result is similar to the finding of Fratzscher et al. (2014).

The oil price reacts negatively to a shock in VIX, and the response lasts about six months. A common perception of the relationship is that the oil price tends to fall when VIX is high. Arguably, it is due the fact that during high volatility, investors are prone to invest in less volatile assets, such as U.S. Treasury bonds. This is likely to strengthen the dollar and have a negative effect on the oil price. The Baa-Aaa credit spread shows a similar effect on the oil price. The shock has a negative impact on the price of oil, which seems to diminish after two months. In total, the shock lasts about 5 months.

Moreover, the oil price initially reacts negatively to a shock in S&P 500. After three months, the shock has a positive effect on the oil price, but is nullified after about five months. Previous findings on this relationship show inconclusive results. Sadorsky (1999) finds a positive relation, while Balcilar et al. (2015) and Zhu et al. (2017) report a negative effect. The results seem to depend on the time period in which the studies examine the effects.

Lastly, lagged values of the oil price shows a positive response to a shock in the price of oil. As expected, lagged values of the oil price can also serve as a good indicator of the oil price. This is not given any attention further in our study, as our focus is on the economic indicators.

Figure 5.1: Impulse Response of Oil Price to Shocks of Economic Factors



Impulse responses (in %) of oil price (Oil) to one standard error shock in BaaAaa (Baa-Aaa), ExcRate (Exchange rate), IndPro (Industrial Production), IntRate (Interest rate), Oil (Oil price), VIX (VIX) and SP500 (S&P500), computed from the VAR model (dashed lines: 95% confidence intervals). Sample period: 1986M01-2017M06.

5.5 Breakpoint Test

Before running the Markov regime switching model, the Chow test is evaluated to determine whether the time series contain any structural breaks. Table 5.4 reports the results of the Chow test. Since the F-statistic is grater than the associated probability, the null hypothesis of no structural breaks present is rejected. As seen from the time series plot in figure 4.1, crude oil seems to have two structural breaks around 1990 and 2008. The 1990 structural break can be linked to the Iraq conflict, and the latter, the global financial crisis, similar to the findings of Salisu and Fasanya (2013).

Table 5.4: Chow Breakpoint Test

F-statistic	2.451141	Prob. F(18,341)	0.0010
Log likelihood ratio	45.87105	Prob. Chi-Square(18)	0.0003
Wald Statistic	44.12053	Prob. Chi-Square(18)	0.0006

Chow test for the evaluation of structural breaks for the seven time series in first differences with the initial breakpoint set to 2008M06. The test is performed on the sample period of 1986M01-2017M06.

5.6 Markov Regime Switching Model

We proceed with the modeling of the Markov regime switching model and apply all variables in first differences, with oil as the dependent variable and Baa-Aaa, industrial production, interest rate, exchange rate, S&P500 and VIX as the independent variables. As described in the methodology section, we also account for seasonality effects in the time series. The top portion of table 5.5 displays the regime coefficients of the variables. The middle section reports the transition probabilities, which denote the probability of switching from one regime to another. Here, regime 1 and 2 denotes the high- and low volatility regime, respectively. The bottom section reports the expected durations, 2.3 months in regime 1 and 19.3 months in regime 2. This indicates that the high volatility regime fades out more quickly than the low volatility regime, implying that the high volatility shocks may not be as persistent. This might also indicate that the volatility of the oil price spends more time in the low volatility regime, being more stable than in the high volatility regime, similar to the findings of Choi and Hammoudeh (2010). A more detailed description of the table with seasonality effects and Autoregressive (AR) terms can be seen in table B.4 and table B.5 in Appendix B.

Table 5.5: Markov Regime Switching Coefficients

Variable	Regime 1	Prob	Regime 2	Prob
C	-0.199752	0.0000	0.027346	0.0420
BaaAaa	-0.055377	0.6036	-0.130741	0.0255
IndPro	4.890063	0.0000	0.271631	0.5543
IntRate	0.129576	0.0187	0.041284	0.5199
SP500	-0.202205	0.3869	-0.306427	0.0055
ExcRate	-0.102141	0.8650	-1.685283	0.0000
VIX	-0.265981	0.0002	-0.062959	0.0452
Transition Probabilities				
Regime 1	0.568432		0.431568	
Regime 2	0.051943		0.948057	
Expected Durations				
	2.317131		19.25202	

First section: Regime coefficients of the Markov regime switching model performed on all seven variables in first differences, with its corresponding probabilities (Prob) for Regime 1 and 2. Transition probabilities for each regime stated in the second section and duration probabilities for each regime in the last section. The variables are denoted as BaaAaa (Baa-Aaa), IndPro (industrial production), IntRate (interest rate), SP500 (S&P 500), ExcRate (exchange rate) and VIX. Sample period is 1986M01-2017M06.

The results from table 5.5 demonstrate that there exist low- and high volatility regimes. As can be seen in the table, the coefficients of the variables vary between the regimes.

The coefficients of the Baa-Aaa credit spread are negative in both regimes. Its impact on the oil price is significant in regime 2, defined as the low volatility regime. This indicates that as the credit spread widens, the oil price decreases. This is an interesting finding since few papers to our knowledge has analyzed the impact of the Baa-Aaa spread on the oil price. However, Gertler and Lown (1999) argue that high-yield bond spreads can be one of the most significant indicators that describe the fluctuations of the oil price, as it reflects the willing to take on risk.

The coefficient of industrial production is positive and significant in regime 1. In regime 2, it has a smaller impact on the oil price, but the effect is still positive. In other words, the findings suggest that industrial production has an overall positive impact on the oil price, which can be confirmed by similar studies as seen from Hamilton (2003), Hamilton (2005) and Lippi and Nobili (2012).

Even though there have been few studies on the direct relationship between oil price and interest rate, it has been shown that the interest rate has an effect on the oil price. As shown from the results of the Markov regime switching model, the interest rate seems to have a significant effect on the oil price in regime 1. This effect diminishes in the low volatility regime and is no longer statistically significant. In brief, the results indicates that the oil price increases as the interest rate raises. Hamilton and Kim (2000) argue that the interest rate can be used as an indicator of the oil price.

The S&P 500 coefficients are negative in both regimes, thus the oil price is inclined to fall when the S&P 500 index rises. The coefficient probabilities show a significant effect for both regimes, a finding similar to Zhu et al. (2017). Sadorsky (1999) and Hammoudeh et al. (2004) find that stocks have a positive effect on the oil price. Thus, it does not seem to be a consensus on the stock markets effect on the oil price.

This study also shows that the exchange rate has a negative impact on the oil price in the high volatility regime. The effect is also statistically significant. In regime 2, the impact is even stronger, but the effect is not as statistically significant. Overall, this finding is similar to previous studies on the relationship between exchange rate and oil price, such as Pindyck and Rotemberg (1988) and Akram (2009), who find a negative effect from the exchange rate on the oil price.

The results also indicate a negative impact from VIX in both regimes. The effect is statistically significant in regime 2, but not in regime 1. The oil price is inclined to fall as the VIX rises, which is in agreement with the work of Ryan and Whiting (2017), who refer to VIX as a potentially significant indicator for oil price movements.

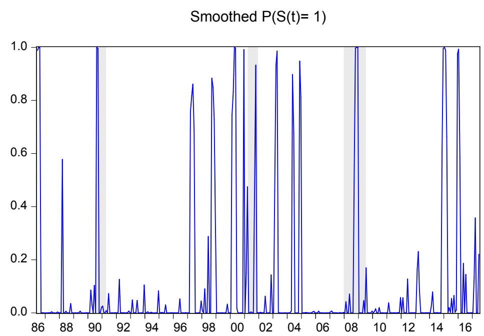
A final note on the seasonality effect is that it is overall more negative in regime 2, however not as statistically significant as in regime 1, as seen in table B.4 in Appendix B. This indicates that the seasonality effects have a larger impact on the oil price during a high volatility regime. The seasonality effects are in general positive in regime 1.

The figures in 5.2 display the filtered and smoothed regime probabilities for both regimes. The filtered probabilities use past information to estimate, whilst the smoothed probabilities use both past and future information. The shaded (grey) areas in the graphs label the NBER recession phases in the US economy. Visually, the graphs suggest that the models have higher smoothed recession probabilities during NBER recessions. All the NBER recessions are well represented by both the filtered- and smoothed probabilities. The high volatility regime, (P(S(t)=1), has smoothed probability of the high volatility regime greater than or equal to 0.5. The low volatility regime, (P(S(t)=2), is characterized by having the high volatility regime lower or equal to 0.5.

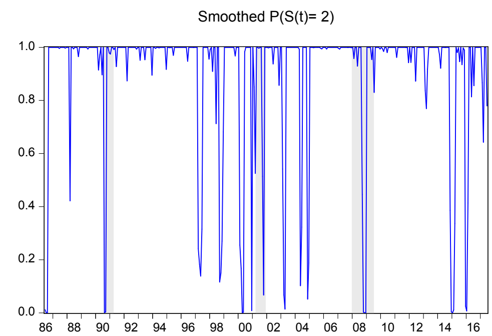
Figures 5.3 and 5.4 show the smoothed probabilities and the timing of change in regime of the high volatility regime. It can be seen that the oil price do not seem to have highly persistent high volatility regime, but instead indicates frequent occurrences of high volatility regime across the whole period, especially in the years of 1997-2005, 2008, and 2014-2016 corresponding to the events 9/11, the US led invasion of Iraq in 2003, oil price spikes in 2008, following the financial crisis. The years of 2014-2016 are characterized of high production in both Russia and United States which caused oil prices to crash.

Figure 5.2: Smoothed and Filtered Regime Probabilities for Regime 1 and Regime 2

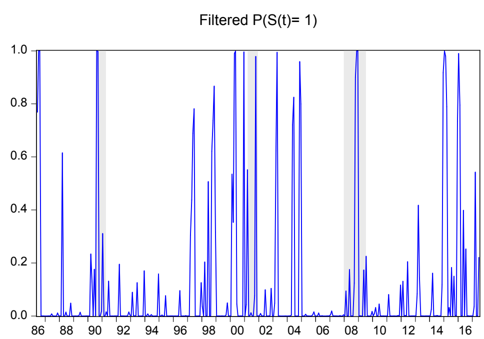
(a) Smoothed Probabilities for Regime 1



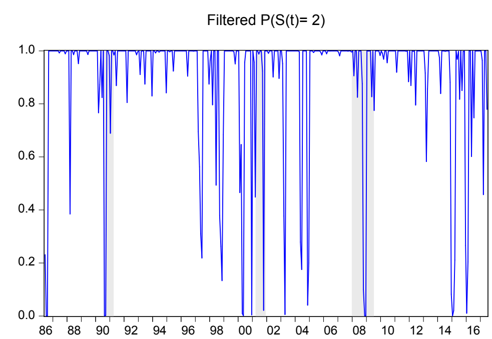
(b) Smoothed Probabilities for Regime 2



(c) Filtered Probabilities for Regime 1

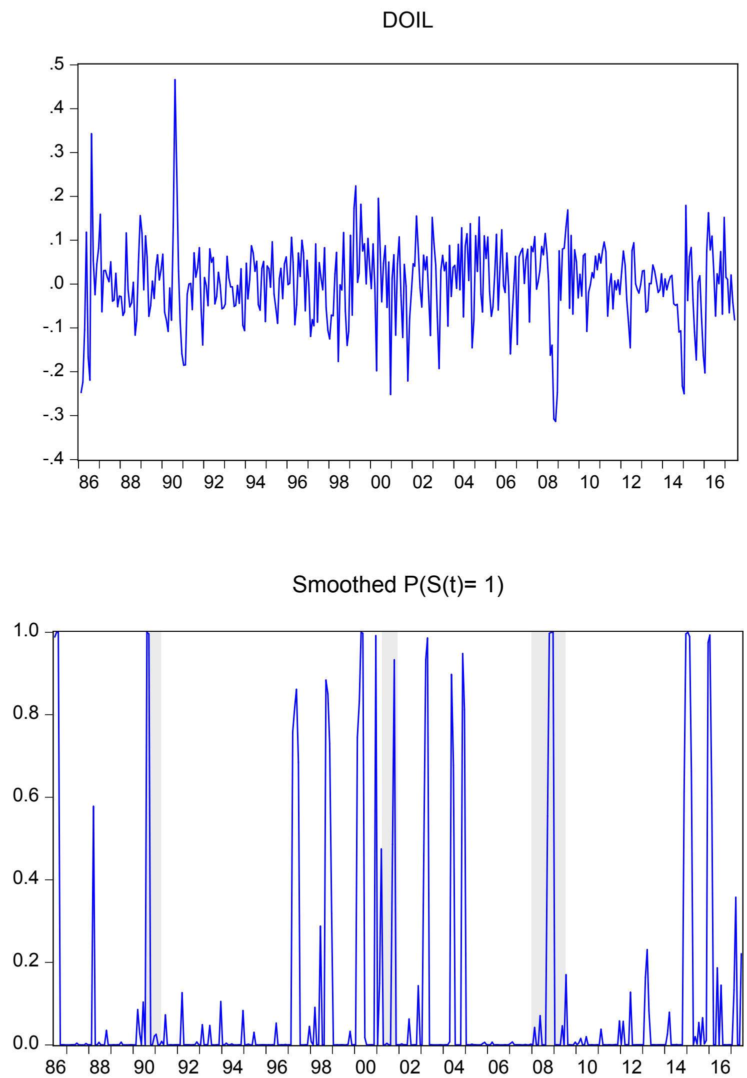


(d) Filtered Probabilities for Regime 2



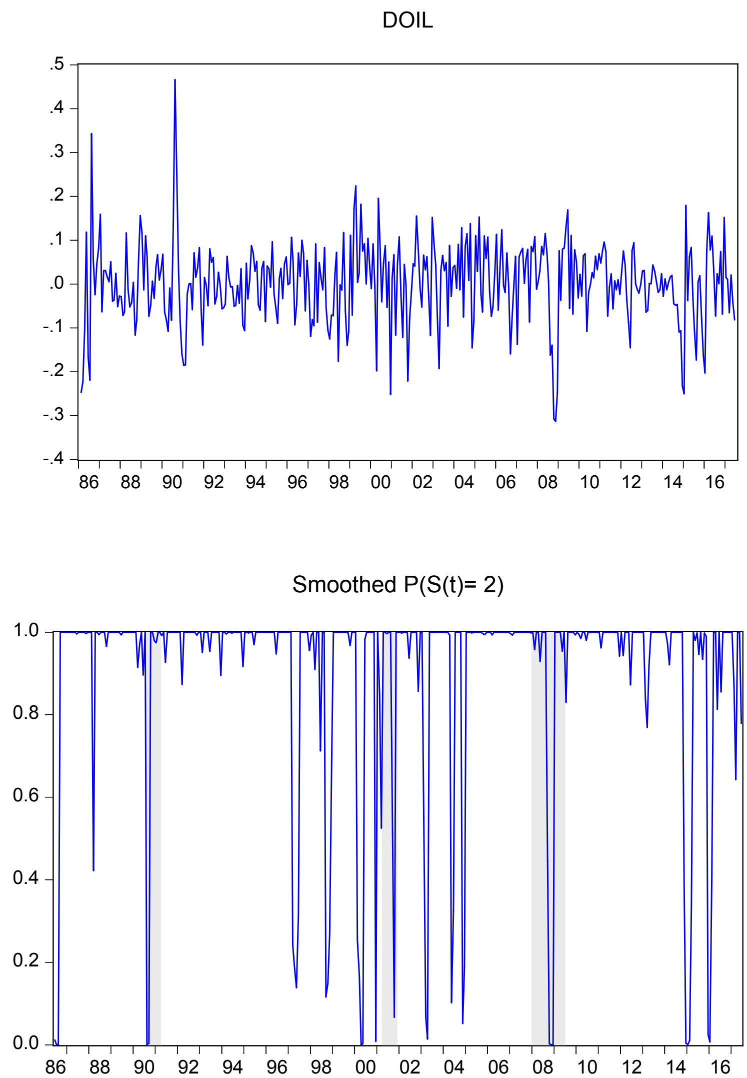
This figures display the smoothed- and filtered regime probabilities of regime 1 and regime 2 in the sample period 1986M01-2017M06. The high volatility regime, regime 1, is denoted by P(S(t)=1, and the low volatility regime, regime 2, is denoted by P(S(t)=2

Figure 5.3: Oil price and Smoothed Probabilities in Regime 1



The graph on top is the actual oil price (in first differences). Underneath are the smoothed probabilities of the first regime. The sample period is 1986M01-2017M06.

Figure 5.4: Oil price and Smoothed Probabilities in Regime 2



The graph on top is the actual oil price (in first differences). Underneath is the smoothed probabilities of the second regime. The sample period is 1986M01-2017M06.

6 CONCLUSION

This study examines the role of six financial- and macroeconomic indicators in explaining the behaviour of the Brent crude oil price using monthly data in the period between January 1986 to June 2017. A multivariate Vector Autoregressive model with Impulse response function and a Markov regime switching regression model are developed in order to analyze the impact. The variables are chosen based on previous findings from the literature and economic theory and include interest rate, exchange rate, industrial production, VIX, Baa-Aaa spread, and S&P 500. The models apply the variables in logarithmic first differences and are adjusted to account for seasonality effects.

One of the most intriguing findings of this study is the negative influence from VIX on the oil price. In particular, the effect seems to last about six months according to the impulse response function. The Markov regime switching analysis confirms this finding and reports statistically significant results. It also indicates that the VIX affects the oil price to a greater extent in the high volatility regime than in the low volatility regime. Our finding is also in agreement with the study of Ryan and Whiting (2017). One possible explanation of the relationship is that VIX reflects the market sentiment, implying that when the sentiment is low, the market experiences high volatility and uncertainty in the oil price. Therefore, the oil price can experience a decline when the VIX rises.

Similar to this finding, the Baa-Aaa spread has a negative impact on the oil price. The impulse response function has a lasting effect of about six months after a shock in the spread. Along the same lines, the Markov Switching model reports a negative effect in both regimes, but only of significance in the low volatility regime, which suggests that Baa-Aaa can be used as an indicator of the oil price movements, as studied by Gertler and Lown (1999).

As expected, an exchange rate shock has a negative accumulated effect on the oil price. The VAR model shows a statistically significant impact lasting up to seven months. This is confirmed by the Markov regime switching model, which finds a negative relation in both the high- and low volatility regime. However, the effect is only statistically significant in the latter. The exchange rate also has the biggest impact on the oil price in the low volatility regime of all the economic factors we have studied. As for the S&P 500 index, the VAR model suggests that the oil price responds negatively to a shock in S&P 500, with a lasting effect up to five months. Taking a closer look at the Markov regime switching analysis, the effect is negative in both regimes, but only statistically significant in the low volatility regime. Previous work on the subject has debated whether there is in fact a positive or negative relationship.

In the case of positive impacts, both models find that industrial production has a positive effect on the oil price. The impulse response function reports that the response lasts up to seven months. In fact, industrial production is reported to have the largest impact on the oil price of all six variables in the high volatility regime, and this effect is also statistically significant. Our finding is consistent with the work of Hamilton (2003, 2005) and Lippi and Nobili (2012).

Considering the shock of interest rate on the oil price, the results are indecisive. The VAR model reports that an increase in the interest rate has a negative impact on the oil price, which lasts up to eight months. The Markov regime switching model, however, finds that the interest rate has a positive effect in both regimes.

On the basis of the model analysis, we conclude that the economic indicators used in this study have varying impact on the oil price. The impulse response of the VAR model gives a good indication of the impact a shock to the economic indicators have on the oil price and its significance. Analyzing the results from the Markov regime switching model, we stress that it is important to also consider the existence of regime shifts in the time series. Our results form the Markov regime switching model show that the effects can vary between regimes.

There are great benefits of operating with two distinct models in addressing several variables' effect on the oil price. Incorporating both VAR and MS in one study provide us with useful insights regarding the significance of the effect of the economic indicators studied in this thesis. Resulting in a more complex and comprehensive analysis.

6.1 Further Research

Some of the shortcomings in this thesis should be considered in further research. Since the VAR model is sensitive to changes in variables, re-evaluating it by exploring different combinations or including additional variables can provide us with interesting results. Another approach would be to split the dataset into shorter time horizons, which can result in a more detailed picture of the relationships. Economic indicators have also shown to have varying effect in net importing- and net exporting countries. Applying data specific to either or, may lead to new, significant findings.

The preliminary tests used in this study suggest that the data is a good fit for a Vector Autoregression approach and indicates that the series have a regime-like structure. Therefore, it could be of interest to apply a Markov Regime Switching Vector Autoregressive (MS-VAR) model, as proposed by Krolzig (1997). This approach estimates a VAR model that incorporates changes in each regime and produces Impulse response functions dependent on each regime. The regime-dependent impulse response function capture a more detailed picture of the time series in certain time period, as they emphasize the effects of the shocks within each regime. A further study can also analyze the regime-dependent impulse responses to identify business cycles and associate the findings with aggregate output. This method is however not as developed and researched as the MS and VAR model and careful analysis should therefore be performed and the dependency of the variables extensively considered before the application of the model.

Moreover, forecasting techniques can be used to determine how well the economic indicators in this study predict the oil price out of sample. Trading in oil future contracts as well as oil spot price require accurate models that predict the oil price movements. Accordingly, accurate prediction of the oil price movements can assist investors, stockbrokers and financial managers in corporate and commercial banks in better oil stock valuations, investment decision and trading. However, for prediction and forecasts, its preferable to research within shorter time horizons.

A APPENDIX - DATA

A.1 Abbreviations used in the Appendix

For the description of the tables and figures from this point forward, note that all seven variables are in logarithmic first differences, sometimes denoted by "D" in front of the variable name. The seven variables are often denoted in the tables as DIND_PRO (industrial production), DINT_RATE (interest rate), DEXC_RATE (exchange rate), DSP500 (S&P500), DVIX (VIX) and DBAA_AAA (BaaAaa). Here, S.E is Standard Errors, Mean dep var is Mean dependent variable and SD dep var is Standard dependent variable. AIC (Akaike information criterion), SIC (Schwartz information criterion), HQC (Hannan-Quinn criterion) and DW stat as Durbin Watson statistic.

A.2 Software Description

In this study we use EViews 10th edition, which is a software used for analysis within econometrics and finance, for forecasting and simulation. The main modules we will use in this study is Univariate Time Series Analysis for Unit Root testing, Advanced Multivariate Analysis for Cointegration testing, Multiple Equation Analysis for VAR and its corresponding Impulse responses and Advanced Single Equation Analysis for Switching Regression.

A.3 Data Sources

U.S. Energy Information Administration (EIA)

”The U.S. Energy Information Administration (EIA) is the statistical and analytical agency within the U.S. Department of Energy. EIA collects, analyzes, and disseminates independent and impartial energy information to promote sound policymaking, efficient markets, and public understanding of energy and its interaction with the economy and the environment.”

URL: eia.gov

Federal Reserve Economic Data (FRED)

"The Research Division of the Federal Reserve Bank of St. Louis is responsible for monitoring the economic and financial literature and produces research in the areas of money and banking, macroeconomics, and international and regional economics. This site offers a wealth of economic data and information to promote economic education and enhance economic research. The widely used database FRED is updated regularly and allows 24/7 access to regional and national financial and economic data."

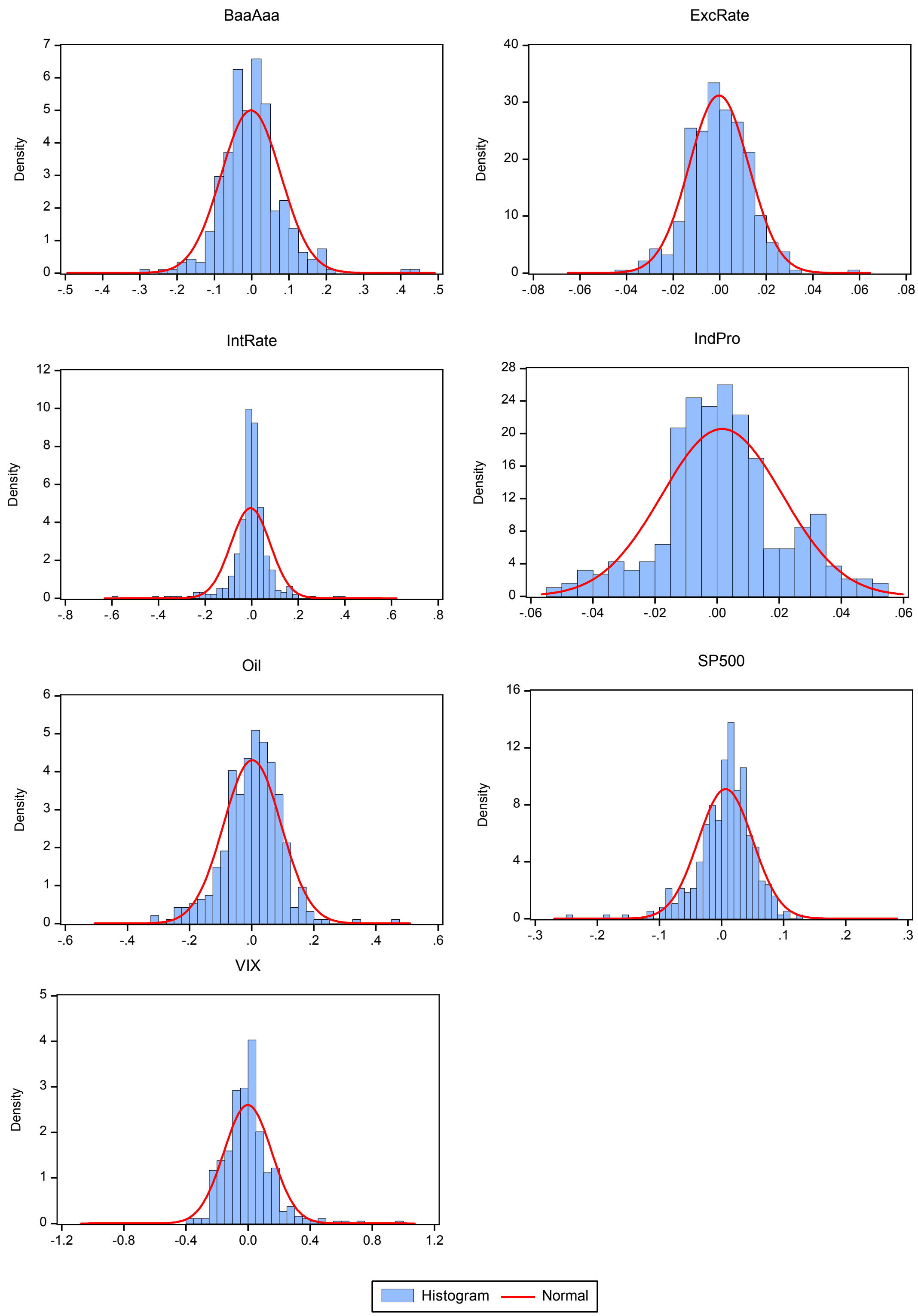
URL: fred.stlouisfed.org

Yahoo Finance

"Yahoo Finance provides financial news, data and commentary including stock quotes, press releases, portoflio management, international market data, financial reports, up-to-date news to manage the financial and economic aspects of your life."

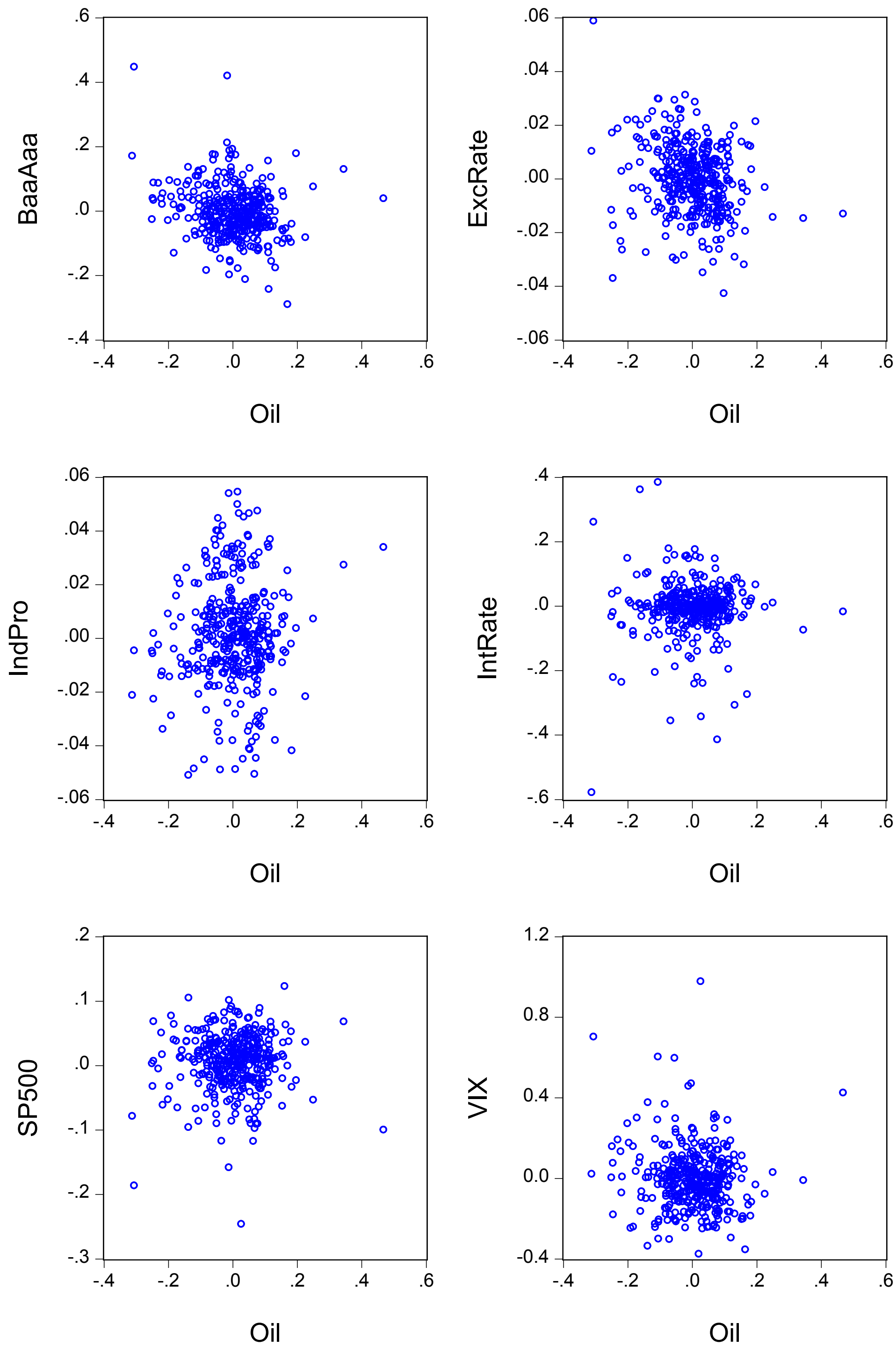
URL: finance.yahoo.com

Figure A.1: Normal Distribution Plots



Normal distribution plots for all seven variables(in first differences) with the corresponding theoretical quantile, time period:1986M01-2016M06, monthly observations. From the top left: BaaAaa (Baa-Aaa), ExcRate (Exchange rate), IntRate (Interest rate), IndPro (Industrial Production), Oil (Oil price), SP500 (S&P500) and VIX (VIX).

Figure A.2: Scatter Plots



Scatter plots for the corresponding six variables against oil in first differences, from the top left: BaaAaa (Baa-Aaa), ExcRate (Exchange rate), IndPro (Industrial Production), IntRate (Interest rate), Oil (Oil price), VIX (VIX) and SP500 (S&P500). Sample period is 1986M01-2017M06.

Table A.1: Unit Root Test in Levels

	ADF	PP	KPSS	DF-GLS
BaaAaa	-4.335915	-3.708005	0.206067	-2.994119
	-3.447534***	-3.447487***	0.739000***	-2.571127***
	-2.869009**	-2.868988**	0.463000**	-1.941668**
	-2.570816*	-2.570805*	0.347000*	-1.616134*
IndPro	-1.696281	-1.265010	2.083436	0.139863
	-3.448062***	-3.447487***	0.739000***	-2.571313***
	-2.869241**	-2.868988**	0.463000**	-1.941694**
	-2.570940*	-2.570805*	0.347000*	-1.616117*
IntRate	-1.441603	-1.693506	1.701988	-0.204795
	-3.447580***	-3.447487***	0.739000***	-2.571143***
	-2.869029**	-2.868988**	0.463000**	-1.941671**
	-2.570827*	-2.570805*	0.347000*	-1.616133*
Oil	-2.027438	-1.348318	1.733544	-1.651508
	-3.447534***	-3.447487***	0.739000***	-2.571127***
	-2.869009**	-2.868988**	0.463000**	-1.941668**
	-2.570816*	-2.570805*	0.347000*	-1.616134*
SP500	0.584018	0.391611	1.958422	2.356282
	-3.447487***	-3.447487***	0.739000***	-2.571110***
	-2.868988**	-2.868988**	0.463000**	-1.941666**
	-2.570805*	-2.570805*	0.347000*	-1.616136*
ExcRate	-2.275639	-2.777280	0.203025	-0.662840
	-3.447580***	-3.447487***	0.739000***	-2.571143***
	-2.869029**	-2.868988**	0.463000**	-1.941671**
	-2.570827*	-2.570805*	0.347000*	-1.616133*
VIX	-5.158870	-5.272481	0.109079	-5.051331
	-3.447487***	-3.447487***	0.739000***	-2.571110***
	-2.968988**	-2.868988**	0.463000**	-1.941666**
	-2.670805*	-2.570805*	0.347000*	-1.616136*

Unit root tests in levels for all seven variables starting BaaAaa (Baa-Aaa), IndPro (industrial production), IntRate (interest rate), SP500 (S&P 500), ExcRate (exchange rate) and VIX. Note: *, **, *** represent significance at 10, 5 and 1 per cent, respectively. The first row for each variable denotes the associated t-statistic. Sample period: 1986M01-2017M06.

Table A.2: Unit Root Tests in First Differences

	ADF	PP	KPSS	DF-GLS
BaaAaa	-13.43209	-13.04003	0.065715	-12.06347
	-3.447534***	-3.447534***	0.739000***	-2.571127***
	-2.869009**	-2.869009**	0.463000**	-1.941668**
	-2.570816*	-2.570816*	0.347000*	-1.616134*
IndPro	-4.647157	-32.01723	0.191469	-2.794096
	-3.448211***	-3.447534***	0.739000***	-2.571313***
	-2.869307**	-2.869009**	0.463000**	-1.941694**
	-2.570975*	-2.570816*	0.347000*	-1.616117*
IntRate	-11.05475	-11.02456	0.086334	-10.96103
	-3.447534***	-3.447534***	0.739000***	-2.571127***
	-2.869009**	-2.869009**	0.463000**	-1.941668**
	-2.570816*	-2.570816*	0.347000*	-1.616134*
Oil	-15.20074	-14.82639	0.072828	-0.740440
	-3.447534***	-3.447534***	0.739000***	-2.571296***
	-2.869009**	-2.869009**	0.463000**	-1.941692**
	-2.570816*	-2.570816*	0.347000*	-1.616119*
SP500	-18.30741	-18.29881	0.133321	-1.951970
	-3.447534***	-3.447534***	0.739000***	-2.571227***
	-2.869009**	-2.869009**	0.463000**	-1.941682**
	-2.570816*	-2.570816*	0.347000*	-1.616125*
ExcRate	-12.61859	-13.12696	0.287933	-0.739105
	-3.447680***	-3.447534***	0.739000***	-2.571278***
	-2.869029**	-2.869009**	0.463000**	-1.941689**
	-2.570827*	-2.570816*	0.347000*	-1.616121*
VIX	-5.158870	-5.272481	0.109079	-5.051331
	-3.447487***	-3.447487***	0.739000***	-2.571110***
	-2.868988**	-2.868988**	0.463000**	-1.941666**
	-2.570805*	-2.570805*	0.347000*	-1.616136*

Unit Root tests in first differences for all seven variables starting BaaAaa (Baa-Aaa), IndPro (industrial production), IntRate (interest rate), SP500 (S&P 500), ExcRate (exchange rate) and VIX. Note: *, **, *** represent significance at 10, 5 and 1 per cent, respectively. The first row for each variable denotes the associated t-statistic. Sample period: 1986M01-2017M06.

Table A.3: Heteroscedasticity Test: White's

Null hypothesis: Homoskedasticity				
F-statistic	2.356267	Prob.F(27,349)	0.0002	
Obs*R-squared	58.12729	Prob.χ^2(24)	0.0005	
Scaled explained SS	123.5732	Prob.χ^2(27)	0.0000	
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.005573	0.001347	4.135866	0.0000
$IndPro^2$	-0.003988	1.436438	-0.002776	0.9978
IndPro*IntRate	-0.448876	0.720656	-0.622871	0.5338
IndPro*ExcRateE	-5.363058	3.803921	-1.409876	0.1595
IndPro*SP500	1.423347	1.190915	1.195171	0.2328
IndPro*VIX	1.233274	0.390654	3.156944	0.0017
IndPro*BaaAaa	0.723158	0.705622	1.024851	0.3061
IndPro	0.017628	0.044258	0.398312	0.6906
$IntRate^2$	0.098199	0.052368	1.875177	0.0616
IntRate*ExcRate	1.465557	1.101776	1.330177	0.1843
IntRate*SP500	-0.146664	0.332124	-0.441594	0.6591
IntRate*VIX	-0.080508	0.087136	-0.923928	0.3562
IntRate*BaaAaa	-0.210684	0.110927	-1.899292	0.0584
IntRate	-0.000536	0.012354	-0.043367	0.9654
$ExcRate^2$	8.679793	4.206039	2.063650	0.0398
ExcRate*SP500	1.140114	1.807140	0.630894	0.5285
ExcRate*VIX	-1.422223	0.551174	-2.580351	0.0103
ExcRate*BaaAaa	-1.026094	0.965393	-1.062878	0.2886
ExcRate	-0.194349	0.071179	-2.730448	0.0066
$SP500^2$	0.228661	0.410575	0.556927	0.5779
SP500*VIX	-0.133864	0.194021	-0.689947	0.4907
SP500*BaaAaa	-0.135525	0.314371	-0.431097	0.6667
SP500	0.012083	0.025613	0.471768	0.6374
VIX^2	-0.043068	0.039212	-1.098344	0.2728
VIX*BaaAaa	0.208817	0.086335	2.418685	0.0161
VIX	0.007552	0.007596	0.994330	0.3208
$BaaAaa^2$	0.025220	0.068365	0.368904	0.7124
BaaAaa	0.024694	0.012287	2.009684	0.0452
R-squared	0.154184	Mean dep var		0.007723
Adjusted R-squared	0.088748	S.D. dep var		0.016248
S.E. of regression	0.015510	AIC		-5.423270
Sum squared resid	0.083958	SIC		-5.131220
Log likelihood	1050.286	HQC.		-5.307348
F-statistic	2.356267	DW stat		1.509643
Prob(F-statistic)	0.000227			

Heteroskedasticity test: White's, for BaaAaa (Baa-Aaa), ExcRate (Exchange rate), IndPro (Industrial Production), IntRate (Interest rate), Oil (Oil price), VIX (VIX) and SP500 (S&P500) in first differences, sample period: 1986M01-2017M01, The null hypothesis assumes homoscedasticity.

Table A.4: Heteroskedasticity Test: Breusch-Pagan-Godfrey

Null hypothesis: Homoskedasticity				
F-statistic	2.733868	Prob.F(6,370)		0.0130
Obs*R-squared	16.00403	Prob.$\chi^2(6)$		0.0137
Scaled explained SS	34.02307	Prob.$\chi^2(6)$		0.0000
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.007663	0.000843	9.094073	0.0000
DIND_PRO	0.004975	0.043780	0.113637	0.9096
DINT_RATE	-0.014521	0.010448	-1.389897	0.1654
DEXC_RATE	-0.159133	0.069020	-2.305611	0.0217
DSP500	-0.000339	0.023118	-0.014685	0.9883
DVIX	0.010467	0.006677	1.567747	0.1178
DBAA_AAA	0.028308	0.010936	2.588588	0.0100
R-squared	0.042451	Mean dep var		0.007723
Adjusted R-squared	0.026923	S.D. dep var		0.016248
S.E. of regression	0.016028	AIC		-5.410601
Sum squared resid	0.095048	SIC		-5.337589
Log likelihood	1026.898	HQC		-5.381621
F-statistic	2.733868	DW stat		1.462391
Prob(F-statistic)	0.013041			

Heteroskedasticity test: Breusch-Pagan-Godfrey of BaaAaa (Baa-Aaa), ExcRate (Exchange rate), IndPro (Industrial Production), IntRate (Interest rate), Oil (Oil price), VIX (VIX) and SP500 (S&P500). The sample period is 1986M01-2017M06. Null hypothesis assumes homoscedasticity.

Table A.5: Least-Squares Residuals

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.002314	0.004664	0.496147	0.6201
BaaAaa	-0.211566	0.060526	-3.495443	0.0005
IndPro	0.338971	0.242311	1.398913	0.1627
IntRate	0.105259	0.057826	1.820275	0.0695
SP500	-0.209442	0.127951	-1.636887	0.1025
ExcRate	-1.635045	0.382005	-4.280172	0.0000
VIX	-0.059153	0.036953	-1.600755	0.1103
R-squared	0.099074	Mean dep var		0.001986
Adj R-squared	0.084465	SD dep var		0.092711
SE of regression	0.088709	AIC		-1.988520
Sum squared resid	2.911629	SIC		-1.915508
Log likelihood	381.8360	HQC.		-1.959540
F-statistic	6.781463	DW stat		1.662824
Prob(F-stat)	0.000001			

Table A.6: HAC (Newey-White) Correction

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.002314	0.005077	0.455778	0.6488
BaaAaa	-0.211566	0.080418	-2.630841	0.0089
IndPro	0.338971	0.254601	1.331382	0.1839
IntRate	0.105259	0.077031	1.366451	0.1726
SP500	-0.209442	0.156611	-1.337338	0.1819
ExcRate	-1.635045	0.554946	-2.946312	0.0034
VIX	-0.059153	0.039474	-1.498526	0.1348
R-squared	0.099074	Mean dep var		0.001986
Adj R-squared	0.084465	SD depvar		0.092711
SE of regression	0.088709	AIC		-1.988520
Sum squared resid	2.911629	SIC		-1.915508
Log likelihood	381.8360	HQC		-1.959540
F-stat	6.781463	DW stat		1.662824
Prob(F-stat)	0.000001	Wald F-stat		3.879048
Prob(Wald F-stat)	0.000906			

Table A.5 with the initial residuals standard errors. Table A.6: HAC (Newey-White) correction for heteroscedasticity. BaaAaa (Baa-Aaa), IndPro (industrial production), IntRate (interest rate), SP500 (S&P 500), ExcRate (exchange rate) and VIX in first differences. Period sample: 1986M01-2017M06.

B APPENDIX B RESULTS

B.1 Detailed Results

The equations displayed below are the VAR estimated equations for Oil, Indpro (industrial production), IntRate (interest rate), ExcRate (exchange rate), SP500 (S&P 500), VIX (VIX), BaaAaa (Baa-Aaa credit spread). @MONTH denotes the monthly added seasonality effect at the February (@MONTH=2) to December (@MONTH=12).

$$\begin{aligned}
\mathsf{Oil} = {} & 126.0228 * \mathsf{Oil}(-1) + 3.394237 * \mathsf{IndPro}(-1) + -1.655800 * \mathsf{IntRate}(-1) \\
& + (-3.630180) * \mathsf{ExcRate}(-1) + (-0.594184) * \mathsf{SP500}(-1) + (-23.61662) * \mathsf{VIX}(-1) \\
& + (-24.11805) * \mathsf{BaaAaa}(-1) + (-303,0709) \\
& + \mathsf{C}(9) * (@\mathsf{MONTH} = 2) + \mathsf{C}(10) * (@\mathsf{MONTH} = 3) + \mathsf{C}(11) * (@\mathsf{MONTH} = 4) \\
& + \mathsf{C}(12) * (@\mathsf{MONTH} = 5) + \mathsf{C}(13) * (@\mathsf{MONTH} = 6) + \mathsf{C}(14) * (@\mathsf{MONTH} = 7) \\
& + \mathsf{C}(15) * (@\mathsf{MONTH} = 8) + \mathsf{C}(16) * (@\mathsf{MONTH} = 9) + \mathsf{C}(17) * (@\mathsf{MONTH} = 10) \\
& + \mathsf{C}(18) * (@\mathsf{MONTH} = 11) + \mathsf{C}(19) * (@\mathsf{MONTH} = 12)
\end{aligned} \tag{B.1}$$

$$\begin{aligned}
\mathsf{IndPro} = {} & (-0.219652) * \mathsf{Oil}(-1) + 0.129652 * \mathsf{IndPro}(-1) + (-0.041301) * \mathsf{IntRate}(-1) \\
& + 0.016927 * \mathsf{ExcRate}(-1) + (-0.060499) * \mathsf{SP500}(-1) + 0.153519 * \mathsf{VIX}(-1) \\
& + 0.109269 * \mathsf{BaaAaa}(-1) + (-1.230903) + \mathsf{C}(28) * (@\mathsf{MONTH} = 2) \\
& + \mathsf{C}(29) * (@\mathsf{MONTH} = 3) + \mathsf{C}(30) * (@\mathsf{MONTH} = 4) \\
& + \mathsf{C}(31) * (@\mathsf{MONTH} = 5) + \mathsf{C}(32) * (@\mathsf{MONTH} = 6) \\
& + \mathsf{C}(33) * (@\mathsf{MONTH} = 7) + \mathsf{C}(34) * (@\mathsf{MONTH} = 8) \\
& + \mathsf{C}(35) * (@\mathsf{MONTH} = 9) + \mathsf{C}(36) * (@\mathsf{MONTH} = 10) \\
& + \mathsf{C}(37) * (@\mathsf{MONTH} = 11) + \mathsf{C}(38) * (@\mathsf{MONTH} = 12)
\end{aligned} \tag{B.2}$$

$$
\begin{aligned}
\mathsf{IntRate} = {} & (-0.553427) * \mathsf{Oil}(-1) + (-0.038172) * \mathsf{IndPro}(-1) + (-2.701757) * \mathsf{IntRate}(-1) \\
& (-0.096749) * \mathsf{ExcRate}(-1) + (-0.009203) * \mathsf{SP500}(-1) + (-0.047321) * \mathsf{VIX}(-1) \\
& + 0.161454 * \mathsf{BaaAaa}(-1) + 3.696741 + \mathsf{C}(47) * (@\mathsf{MONTH} = 2) \\
& + \mathsf{C}(48) * (@\mathsf{MONTH} = 3) + \mathsf{C}(49) * (@\mathsf{MONTH} = 4) \\
& + \mathsf{C}(50) * (@\mathsf{MONTH} = 5) + \mathsf{C}(51) * (@\mathsf{MONTH} = 6) \\
& + \mathsf{C}(52) * (@\mathsf{MONTH} = 7) + \mathsf{C}(53) * (@\mathsf{MONTH} = 8) \\
& + \mathsf{C}(54) * (@\mathsf{MONTH} = 9) + \mathsf{C}(55) * (@\mathsf{MONTH} = 10) \\
& + \mathsf{C}(56) * (@\mathsf{MONTH} = 11) + \mathsf{C}(57) * (@\mathsf{MONTH} = 12)
\end{aligned}
\tag{B.3}
$$

$$
\begin{aligned}
\mathsf{ExcRate} = {} & 0.070874 * \mathsf{Oil}(-1) + 0.010302 * \mathsf{IndPro}(-1) + (-0.108839) * \mathsf{IntRate}(-1) \\
& + (-0.043291) * \mathsf{ExcRate}(-1) + 0.024130 * \mathsf{SP500}(-1) + (-0.0745543) * \mathsf{VIX}(-1) \\
& + (-0.060489) * \mathsf{BaaAaa}(-1) + 0.163043 + \mathsf{C}(66) * (@\mathsf{MONTH} = 2) \\
& + \mathsf{C}(67) * (@\mathsf{MONTH} = 3) + \mathsf{C}(68) * (@\mathsf{MONTH} = 4) \\
& + \mathsf{C}(69) * (@\mathsf{MONTH} = 5) + \mathsf{C}(70) * (@\mathsf{MONTH} = 6) \\
& + \mathsf{C}(71) * (@\mathsf{MONTH} = 7) + \mathsf{C}(72) * (@\mathsf{MONTH} = 8) \\
& + \mathsf{C}(73) * (@\mathsf{MONTH} = 9) + \mathsf{C}(74) * (@\mathsf{MONTH} = 10) \\
& + \mathsf{C}(75) * (@\mathsf{MONTH} = 11) + \mathsf{C}(76) * (@\mathsf{MONTH} = 12)
\end{aligned}
\tag{B.4}
$$

$$
\begin{aligned}
\mathsf{SP500} = {} & 0.019732 * \mathsf{Oil}(-1) + (-0.23163) * \mathsf{IndPro}(-1) + 0.166616\mathsf{IntRate}(-1) \\
& + 0.026451 * \mathsf{ExcRate}(-1) + (-0.113695) * \mathsf{SP500}(-1) + 0.445235 * \mathsf{VIX}(-1) \\
& + 0.039459 * \mathsf{BaaAaa}(-1) + (-4.736667) + \mathsf{C}(85) * (@\mathsf{MONTH} = 2) \\
& + \mathsf{C}(86) * (@\mathsf{MONTH} = 3) + \mathsf{C}(87) * (@\mathsf{MONTH} = 4) \\
& + \mathsf{C}(88) * (@\mathsf{MONTH} = 5) + \mathsf{C}(89) * (@\mathsf{MONTH} = 6) \\
& + \mathsf{C}(90) * (@\mathsf{MONTH} = 7) + \mathsf{C}(91) * (@\mathsf{MONTH} = 8) \\
& + \mathsf{C}(92) * (@\mathsf{MONTH} = 9) + \mathsf{C}(93) * (@\mathsf{MONTH} = 10) \\
& + \mathsf{C}(94) * (@\mathsf{MONTH} = 11) + \mathsf{C}(95) * (@\mathsf{MONTH} = 12)
\end{aligned}
\tag{B.5}
$$

$$
\begin{aligned}
\mathsf{VIX} = {} & (-0.0607780) * \mathsf{Oil}(-1) + 0.083500 * \mathsf{IndPro}(-1) + (-1.000311) * \mathsf{IntRate}(-1) \\
& + 0.013393 * \mathsf{ExcRate}(-1) + 0.535018 * \mathsf{SP500}(-1) + 0.354588 * \mathsf{VIX}(-1) \\
& + 0.563303 * \mathsf{BaaAaa}(-1) + 1.245384 + \mathsf{C}(104) * (@\mathsf{MONTH} = 2) \\
& + \mathsf{C}(105) * (@\mathsf{MONTH} = 3) + \mathsf{C}(106) * (@\mathsf{MONTH} = 4) \\
& + \mathsf{C}(107) * (@\mathsf{MONTH} = 5) + \mathsf{C}(108) * (@\mathsf{MONTH} = 6) \\
& + \mathsf{C}(109) * (@\mathsf{MONTH} = 7) + \mathsf{C}(110) * (@\mathsf{MONTH} = 8) \\
& + \mathsf{C}(111) * (@\mathsf{MONTH} = 9) + \mathsf{C}(112) * (@\mathsf{MONTH} = 10) \\
& + \mathsf{C}(113) * (@\mathsf{MONTH} = 11) + \mathsf{C}(114) * (@\mathsf{MONTH} = 12)
\end{aligned}
\tag{B.6}
$$

$$
\begin{aligned}
\mathsf{BaaAaa} = {} & 0.921374 * \mathsf{Oil}(-1) + 0.014784 * \mathsf{IndPro}(-1) + (-1.109516) * \mathsf{IntRate}(-1) \\
& + (-0.118414) * \mathsf{ExcRate}(-1) + 0.744786 * \mathsf{SP500}(-1) + (-2.550130) * \mathsf{VIX}(-1) \\
& + (-1.674811) * \mathsf{BaaAaa}(-1) + 1.471414 + \mathsf{C}(123) * (@\mathsf{MONTH} = 2) \\
& + \mathsf{C}(124) * (@\mathsf{MONTH} = 3) + \mathsf{C}(125) * (@\mathsf{MONTH} = 4) \\
& + \mathsf{C}(126) * (@\mathsf{MONTH} = 5) + \mathsf{C}(127) * (@\mathsf{MONTH} = 6) \\
& + \mathsf{C}(128) * (@\mathsf{MONTH} = 7) + \mathsf{C}(129) * (@\mathsf{MONTH} = 8) \\
& + \mathsf{C}(130) * (@\mathsf{MONTH} = 9) + \mathsf{C}(131) * (@\mathsf{MONTH} = 10) \\
& + \mathsf{C}(132) * (@\mathsf{MONTH} = 11) + \mathsf{C}(133) * (@\mathsf{MONTH} = 12)
\end{aligned}
\tag{B.7}
$$

Table B.1: Coefficients (C(1)-C(100)) of the VAR Model Estimates

Coefficient	Sum	Mean	Newton Dir.	Coefficient	Sum	Mean	Newton Dir.
C(1)	126.0228	0.335167	39.50769	C(51)	-0.134896	-0.000359	-0.080030
C(2)	3.394237	0.009027	4.155555	C(52)	0.523267	0.001392	-0.071977
C(3)	-1.655800	-0.004404	-1.240589	C(53)	0.530411	0.001411	-0.082378
C(4)	-3.630180	-0.009655	1.100776	C(54)	0.767106	0.002040	-0.071231
C(5)	-0.594184	-0.001580	-0.792628	C(55)	-0.696983	-0.001854	-0.130719
C(6)	-23.61662	-0.062810	-0.806670	C(56)	1.168327	0.003107	-0.053748
C(7)	-24.11805	-0.064144	-0.769403	C(57)	-0.800312	-0.002128	-0.140465
C(8)	-303.0709	-0.806040	-0.745174	C(58)	0.070874	0.000188	-0.004975
C(9)	-33.15260	-0.088172	-0.623529	C(59)	0.010302	2.74E-05	0.326213
C(10)	-46.60654	-0.123954	-0.593524	C(60)	-0.108839	-0.000289	-0.018397
C(11)	-11.79771	-0.031377	-0.274896	C(61)	-0.043291	-0.000115	-0.754058
C(12)	-46.72804	-0.124277	-1.554755	C(62)	0.024130	6.42E-05	-0.020828
C(13)	9.070531	0.024124	0.078960	C(63)	-0.074543	-0.000198	-0.002730
C(14)	-8.476733	-0.022545	0.970660	C(64)	-0.060489	-0.000161	0.000211
C(15)	-18.28803	-0.048638	-0.068284	C(65)	0.163043	0.000434	-0.006531
C(16)	19.12442	0.050863	0.058360	C(66)	-0.070014	-0.000186	0.006041
C(17)	-13.93866	-0.037071	0.031128	C(67)	-0.139827	-0.000372	-0.000405
C(18)	-30.57778	-0.081324	0.078480	C(68)	0.352422	0.000937	0.017762
C(19)	-62.85748	-0.167174	0.056227	C(69)	0.069885	0.000186	0.009625
C(20)	-0.219652	-0.000584	-0.045445	C(70)	-0.272307	-0.000724	-0.004384
C(21)	0.129761	0.000345	0.096705	C(71)	0.126922	0.000338	0.004596
C(22)	-0.041301	-0.000110	-0.031068	C(72)	-0.009623	-2.56E-05	0.015353
C(23)	0.016927	4.50E-05	-0.069806	C(73)	0.036514	9.71E-05	-0.003588
C(24)	-0.060499	-0.000161	-0.058732	C(74)	0.190468	0.000507	0.012984
C(25)	0.153519	0.000408	-0.015906	C(75)	-0.018409	-4.90E-05	0.005830
C(26)	0.109269	0.000291	0.020989	C(76)	0.085565	0.000228	0.011774
C(27)	-1.230903	-0.003274	-0.001377	C(77)	0.019732	5.25E-05	0.016071
C(28)	-0.472333	-0.001256	-0.013549	C(78)	-0.023163	-6.16E-05	-0.954243
C(29)	-0.373058	-0.000992	-0.010329	C(79)	0.166616	0.000443	0.048735
C(30)	0.927544	0.002467	0.031722	C(80)	0.026451	7.03E-05	0.353820
C(31)	-0.385939	-0.001026	-0.008082	C(81)	-0.113695	-0.000302	0.009570
C(32)	-1.878205	-0.004995	-0.056254	C(82)	0.445235	0.001184	0.045727
C(33)	2.034979	0.005412	0.063667	C(83)	0.039459	0.000105	-0.026468
C(34)	-2.178914	-0.005795	-0.065892	C(84)	-4.736667	-0.012598	-0.008660
C(35)	0.190204	0.000506	0.006269	C(85)	-0.269937	-0.000718	-0.000478
C(36)	0.319540	0.000850	0.011973	C(86)	-0.889005	-0.002364	-0.012548
C(37)	0.520197	0.001384	0.019617	C(87)	-0.906025	-0.002410	-0.015275
C(38)	0.062648	0.000167	0.001263	C(88)	-0.735751	-0.001957	-0.025721
C(39)	-0.553427	-0.001472	-0.121898	C(89)	0.103016	0.000274	0.018085
C(40)	-0.038172	-0.000102	-0.113159	C(90)	-0.540683	-0.001438	0.017652
C(41)	-2.701757	-0.007186	-1.124633	C(91)	0.443325	0.001179	-0.005702
C(42)	-0.096749	-0.000257	-0.410857	C(92)	0.480328	0.001277	0.053882
C(43)	-0.009203	-2.45E-05	-0.108335	C(93)	-0.418562	-0.001113	-0.008664
C(44)	-0.047321	-0.000126	0.085682	C(94)	-0.610039	-0.001622	-0.018183
C(45)	0.161454	0.000429	0.219229	C(95)	-1.039332	-0.002764	-0.028310
C(46)	3.696741	0.009832	0.089765	C(96)	-0.607780	-0.001616	0.000569
C(47)	0.710059	0.001888	-0.102398	C(97)	0.083500	0.000222	0.699423
C(48)	-0.477261	-0.001269	-0.107893	C(98)	-1.000311	-0.002660	-0.560499
C(49)	0.022101	5.88E-05	-0.072311	C(99)	0.013393	3.56E-05	1.081246
C(50)	-0.315010	-0.000838	-0.096044	C(100)	0.535018	0.001423	1.140007

The table reports the coefficients for the estimates of the VAR model from coefficient 1 (C(1)) to number 100 (C(100)). Accounted for lag selection and seasonality. Sample period 1986M01-2017M06.

Table B.2: Coefficients (C(101)-C(133) of the VAR Model Estimates Continued

Coefficient	Sum	Mean	Newton Dir.
C(101)	0.354588	0.000943	0.263278
C(102)	0.563303	0.001498	0.276668
C(103)	1.245384	0.003312	-0.013691
C(104)	0.096366	0.000256	-0.008962
C(105)	0.855112	0.002274	0.032639
C(106)	1.603686	0.004265	0.052044
C(107)	1.465064	0.003896	0.067615
C(108)	-0.440166	-0.001171	-0.004719
C(109)	2.239219	0.005955	0.057900
C(110)	-4.621728	-0.012292	-0.116328
C(111)	-2.374410	-0.006315	-0.100334
C(112)	-4.776821	-0.012704	-0.146742
C(113)	3.675271	0.009775	0.109920
C(114)	3.966138	0.010548	0.134916
C(115)	0.921374	0.002450	0.168010
C(116)	0.014784	3.93E-05	1.393743
C(117)	-1.109516	-0.002951	-0.268707
C(118)	-0.118414	-0.000315	-0.133552
C(119)	0.744786	0.001981	0.594355
C(120)	-2.550130	-0.006782	-0.091358
C(121)	-1.674811	-0.004454	-0.478191
C(122)	1.471414	0.003913	0.060034
C(123)	2.710174	0.007208	0.000588
C(124)	0.653203	0.001737	-0.075758
C(125)	0.415830	0.001106	-0.069889
C(126)	0.627966	0.001670	-0.038640
C(127)	-0.477420	-0.001270	-0.099688
C(128)	-0.209117	-0.000556	-0.098706
C(129)	0.807505	0.002148	0.000937
C(130)	-0.209465	-0.000557	-0.118285
C(131)	-1.698100	-0.004516	-0.105452
C(132)	-1.369221	-0.003642	-0.077467
C(133)	-1.479951	-0.003936	-0.096064

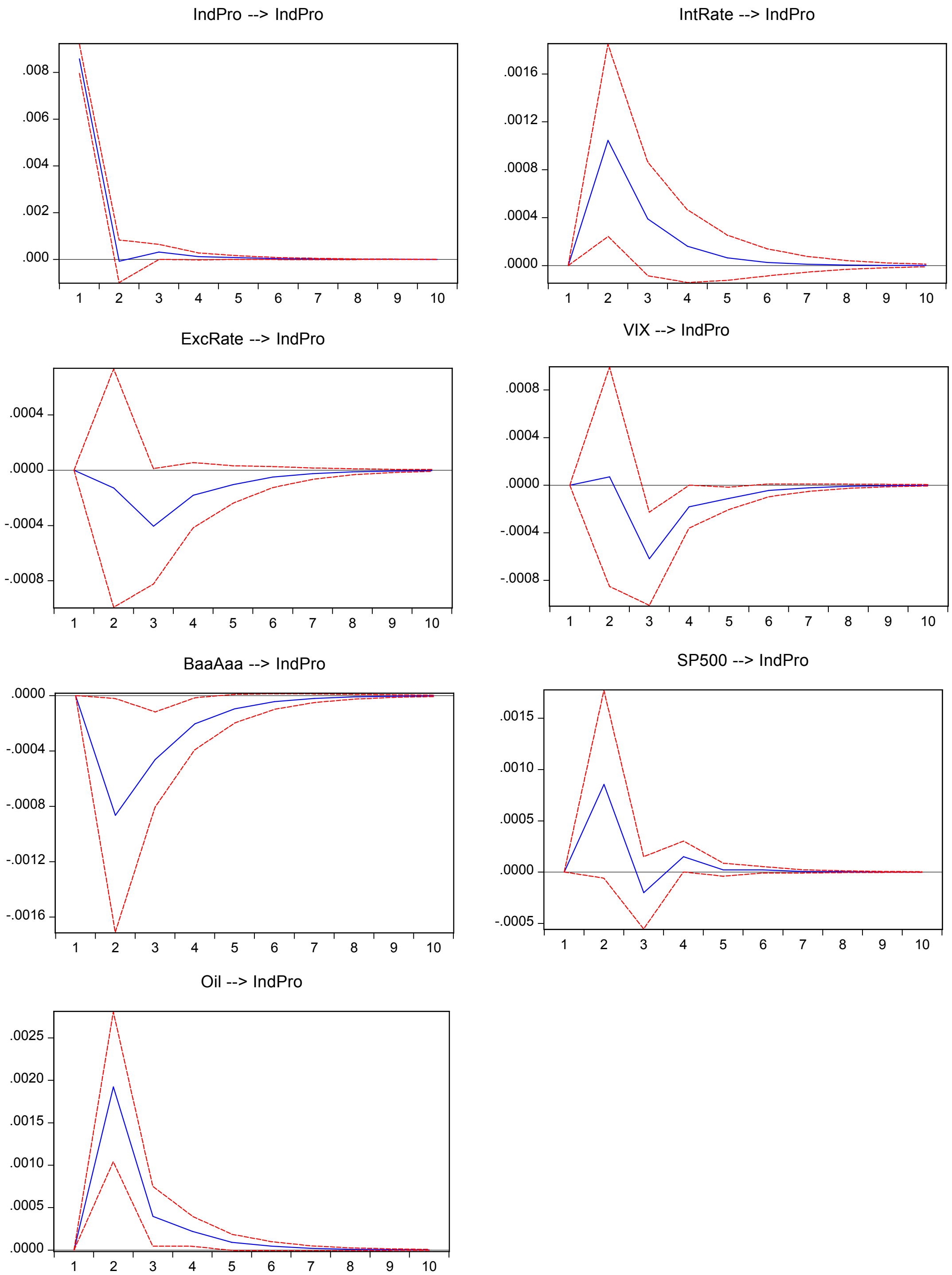
The table reports the coefficients for the estimates of the VAR model from coefficient 101 (C(101)) to 113 (C(113)). Accounted for lag selection and seasonality. Sample period 1986M01-2017M06.

Table B.3: Response of Oil to Shocks to the Economic Variables by Period

Period	Oil	IndPro	IntRate	ExcRate	SP500	VIX	BaaAaa
1	0.087755	0.005245	-0.002112	-0.016837	0.002250	-0.012992	-0.011056
	(0.00320)	(0.00452)	(0.00452)	(0.00448)	(0.00452)	(0.00450)	(0.00451)
2	0.019978	0.006035	-0.002718	-0.015629	0.001757	-0.008420	-0.010722
	(0.00454)	(0.00463)	(0.00408)	(0.00439)	(0.00477)	(0.00473)	(0.00433)
3	0.006452	0.002670	-0.002221	-0.007741	0.003536	-0.003558	-0.005006
	(0.00239)	(0.00153)	(0.00309)	(0.00278)	(0.00205)	(0.00197)	(0.00258)
4	0.002444	0.001351	-0.001453	-0.003680	0.001579	-0.001823	-0.002214
	(0.00127)	(0.00080)	(0.00199)	(0.00154)	(0.00105)	(0.00098)	(0.00132)
5	0.000966	0.000530	-0.000855	-0.001632	0.000691	-0.000675	-0.000873
	(0.00070)	(0.00044)	(0.00119)	(0.00083)	(0.00054)	(0.00051)	(0.00068)
6	0.000365	0.000208	-0.000477	-0.000726	0.000245	-0.000264	-0.000338
	(0.00039)	(0.00023)	(0.00068)	(0.00043)	(0.00029)	(0.00028)	(0.00035)
7	0.000128	7.18E-05	-0.000257	-0.000315	8.41E-05	-8.66E-05	-0.000120
	(0.00021)	(0.00013)	(0.00037)	(0.00022)	(0.00016)	(0.00015)	(0.00018)
8	3.85E-05	2.22E-05	-0.000135	-0.000135	2.24E-05	-2.51E-05	-3.82E-05
	(0.00012)	(6.8E-05)	(0.00020)	(0.00011)	(8.5E-05)	(8.3E-05)	(9.8E-05)
9	7.66E-06	4.62E-06	-6.98E-05	-5.65E-05	2.74E-06	-3.64E-06	-8.94E-06
	(6.3E-05)	(3.7E-05)	(0.00010)	(5.4E-05)	(4.7E-05)	(4.6E-05)	(5.2E-05)
10	-1.41E-06	-5.52E-07	-3.54E-05	-2.30E-05	-2.53E-06	2.03E-06	8.73E-08
	(3.4E-05)	(2.0E-05)	(5.2E-05)	(2.7E-05)	(2.6E-05)	(2.5E-05)	(2.8E-05)

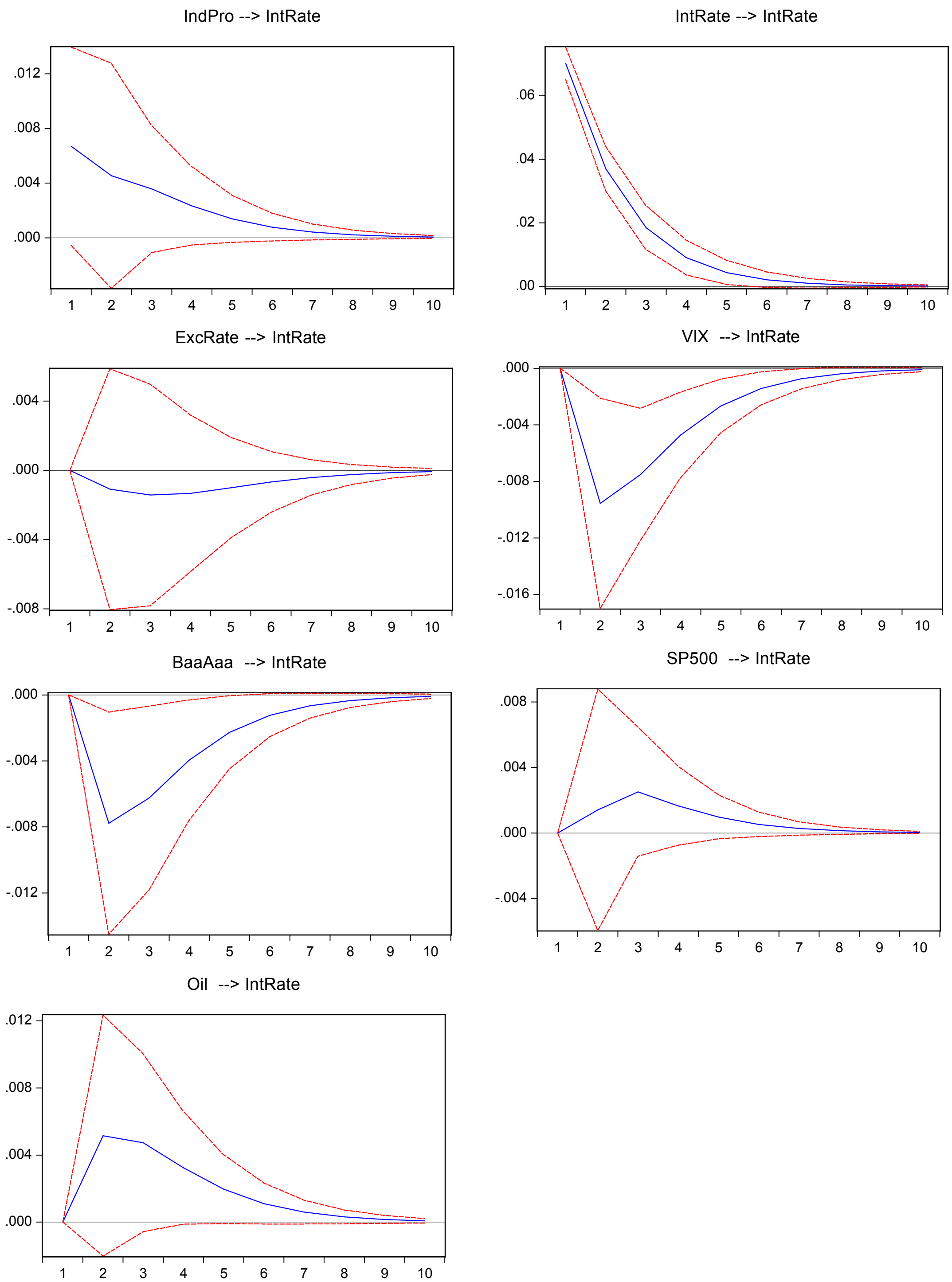
Response of oil to shocks to BaaAaa (Baa-Aaa), IndPro (industrial production), IntRate (interest rate), SP500 (S&P 500), ExcRate (exchange rate), oil price (Oil) and VIX by period (in months). Sample period: 1986M01-2017M06.

Figure B.1: Impulse Response of Industrial Production to Shock in Economic Variables



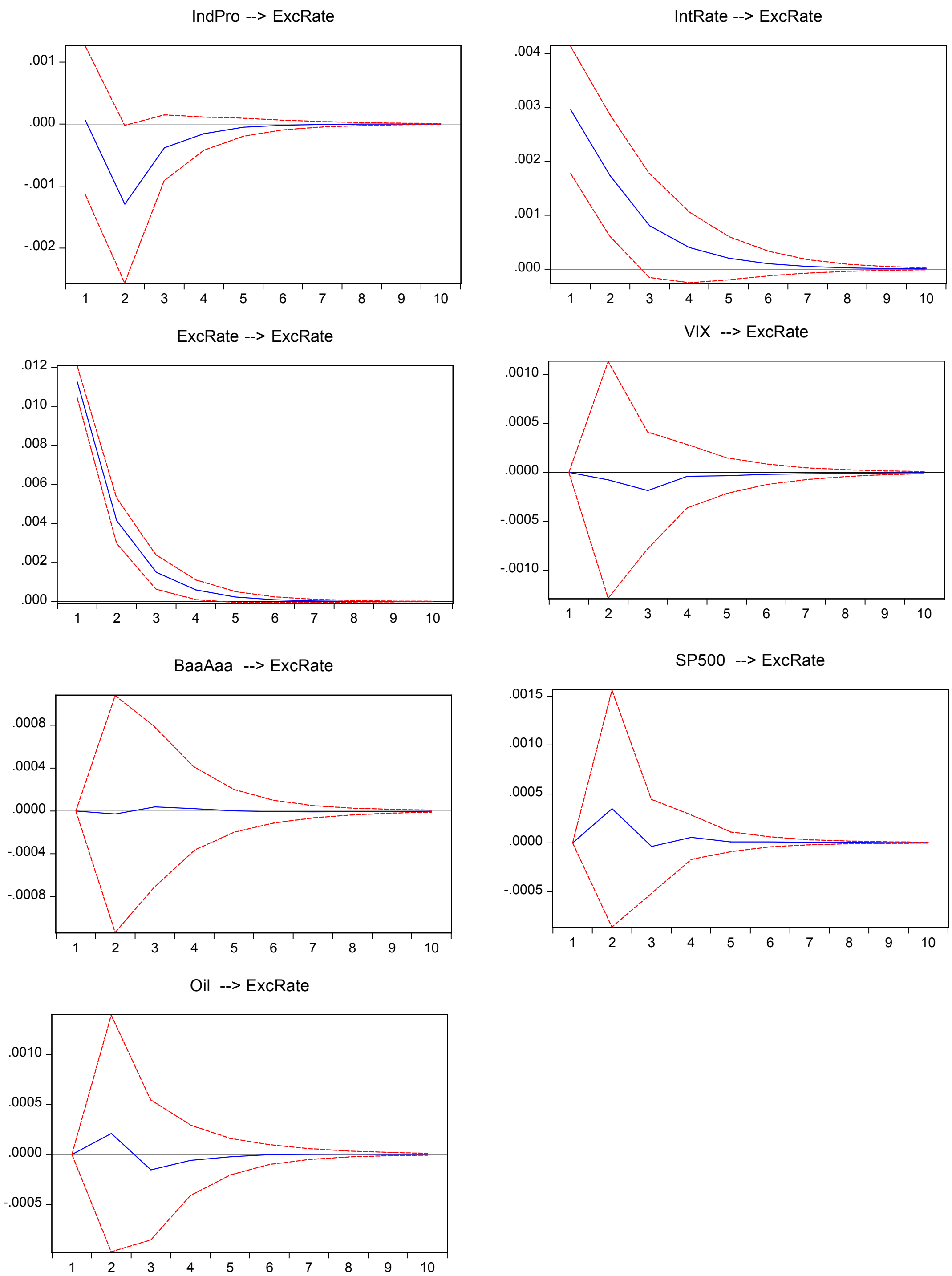
Impulse responses (in %) of BaaAaa (Baa-Aaa), IndPro (industrial production), IntRate (interest rate), SP500 (S&P 500), ExcRate (exchange rate), VIX and Oil computed from the VAR model (dashed lines: 95% confidence intervals). Sample period: 1986M01-2017M06.

Figure B.2: Impulse Response of Real Interest Rate to Shock in Economic Variables



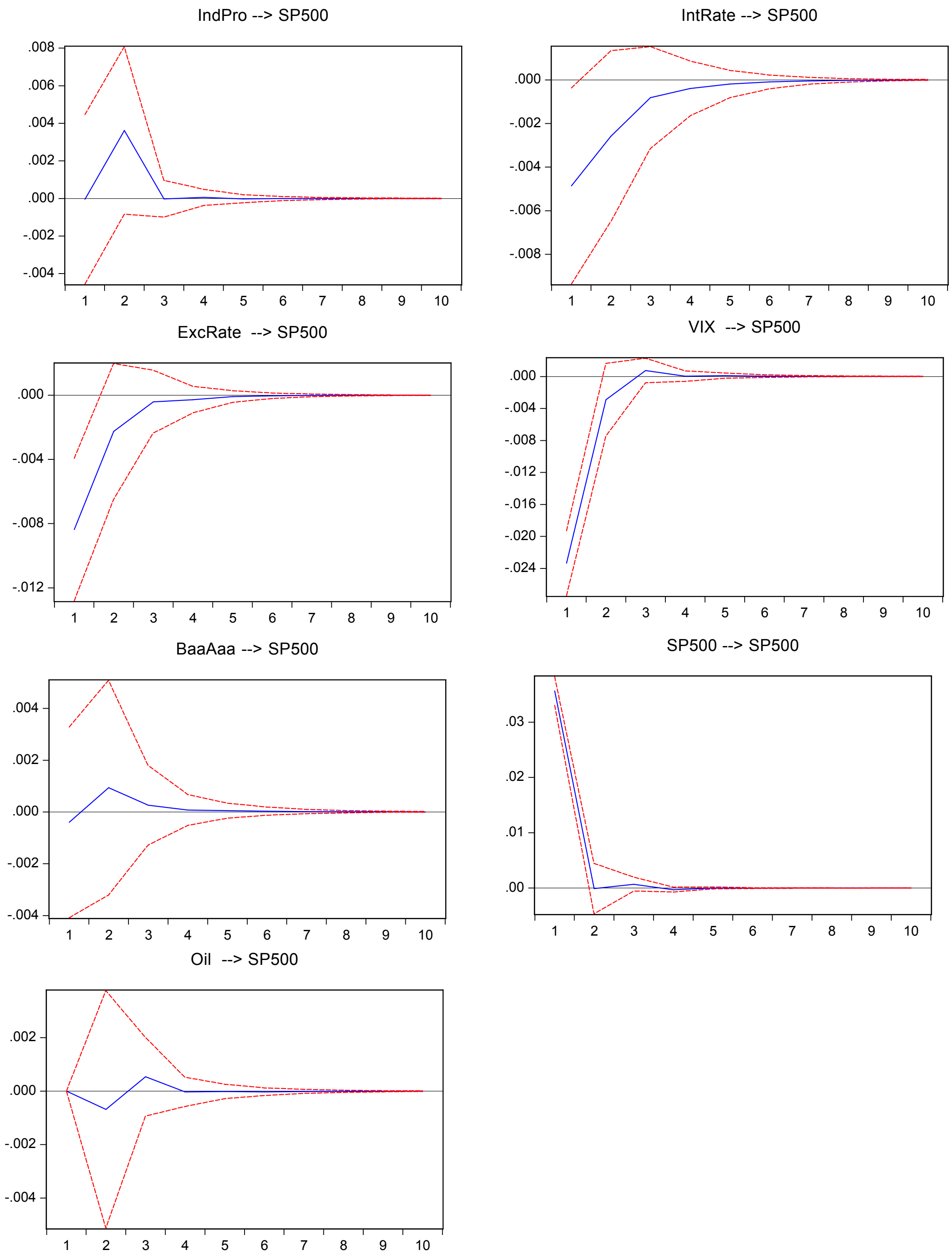
Impulse responses (in %) of IntRate (interest rate) to one standard error shock in IndPro (industrial production), IntRate (interest rate), ExcRate (exchange rate), VIX (VIX), BaaAaa (Baa-Aaa), SP500 (S&P500), and Oil (oil price), computed from the VAR model (dashed lines: 95% confidence intervals). Sample period: 1986M01-2017M06.

Figure B.3: Impulse Response of Exchange Rate to Shock in Economic Variables



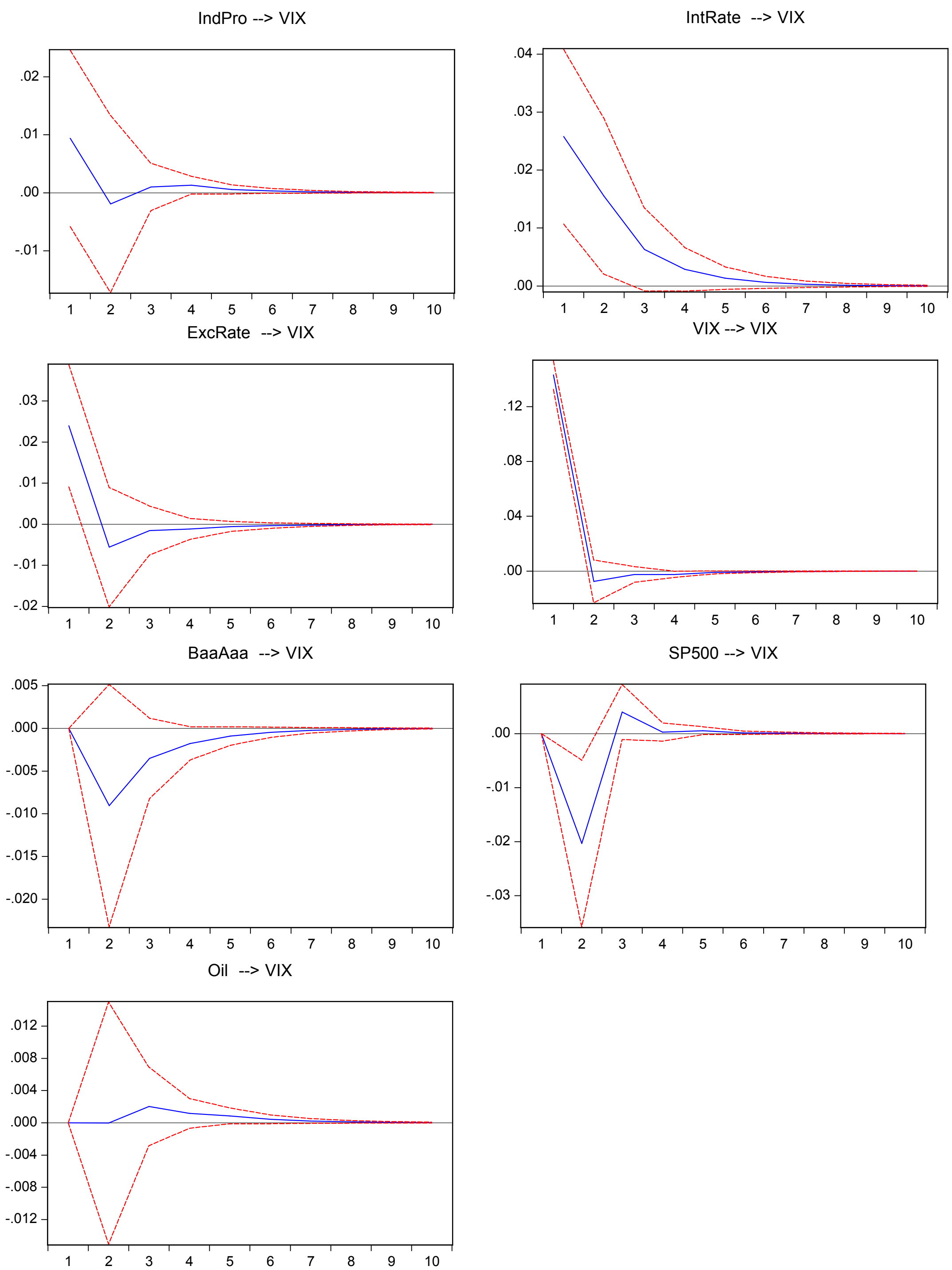
Impulse responses (in %) of ExcRate (exchange rate) to one standard error shock in IndPro (industrial production), IntRate (interest rate), ExcRate (exchange rate), VIX (VIX), BaaAaa (Baa-Aaa), SP500 (S&P500), and Oil (oil price) computed from the VAR model (dashed lines: 95% confidence intervals). Sample period: 1986M01-2017M06.

Figure B.4: Impulse Response of S&P 500 to Shock in Economic Variables



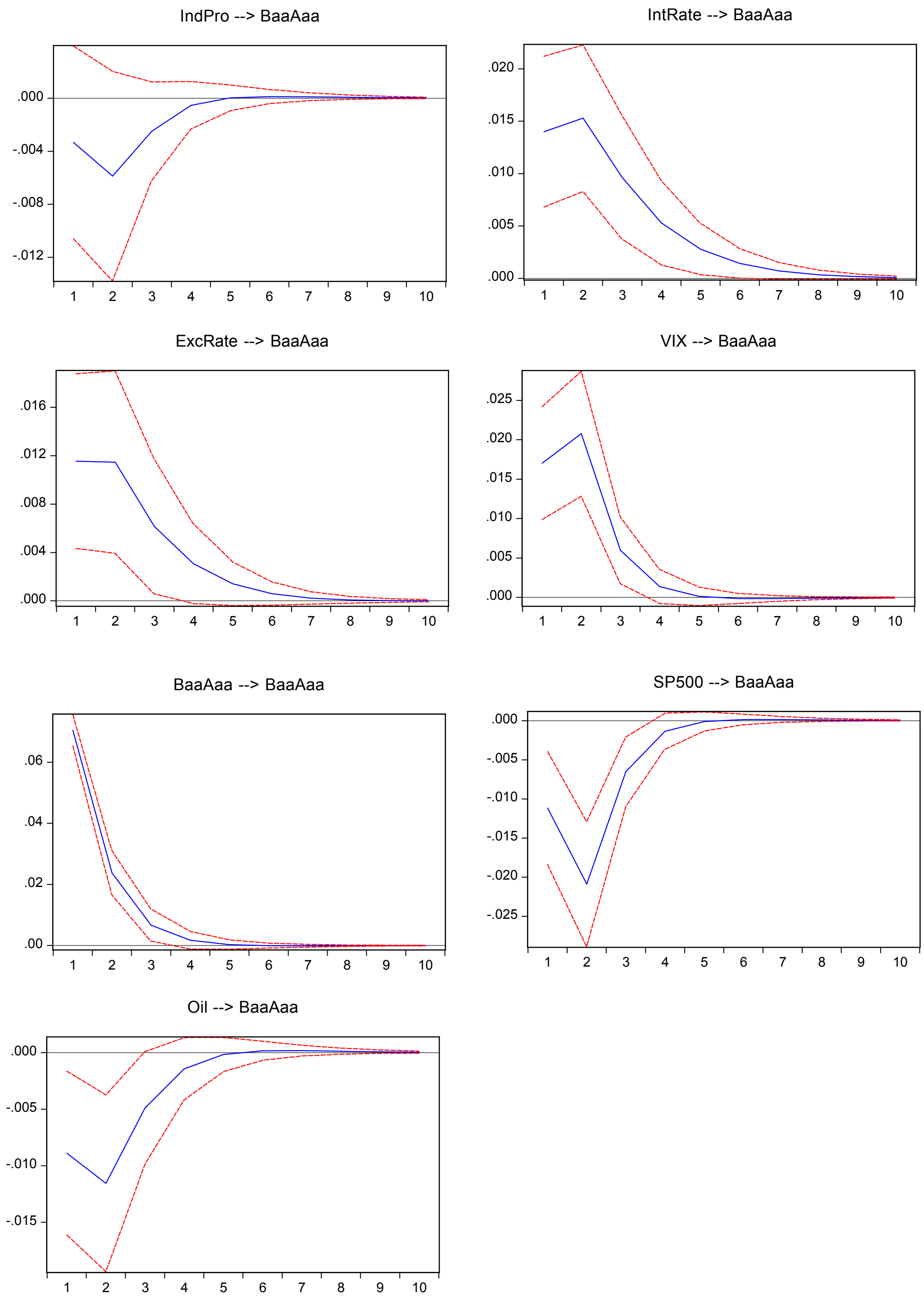
Impulse responses (in %) of SP500 (S&P500) to one standard error shock in IndPro (industrial production), IntRate (interest rate), ExcRate (exchange rate), VIX (VIX), BaaAaa (Baa-Aaa), SP500 (S&P500), and Oil (oil price) computed from the VAR model (dashed lines: 95% confidence intervals). Sample period: 1986M01-2017M06.

Figure B.5: Impulse Response of VIX to Shock in Economic Variables



Impulse responses (in %) of VIX (VIX) to one standard error shock in IndPro (industrial production), IntRate (interest rate), ExcRate (exchange rate), VIX (VIX), BaaAaa (Baa-Aaa), SP500 (S&P500), and Oil (oil price), computed from the VAR model (dashed lines: 95% confidence intervals). Sample period: 1986M01-2017M06.

Figure B.6: Impulse Response of Baa-Aaa to Shock in Economic Variables



Impulse responses (in %) BaaAaa (Baa-Aaa) to one standard error shock in IndPro (industrial production), IntRate (interest rate), ExcRate (exchange rate), VIX (VIX), BaaAaa (Baa-Aaa), SP500 (S&P500), and Oil (oil price), computed from the VAR model (dashed lines: 95% confidence intervals). Sample period: 1986M01-2017M06.

Table B.4: Markov Regime Switching Regression Model

Variable	Coefficient	Std. Error	z-Statistic	Prob.
Regime 1				
C	-0.199752	0.028522	-7.003413	0.0000
BaaAaa	-0.055377	0.106656	-0.519213	0.6036
IndPro	4.899063	0.806285	6.076095	0.0000
IntRate	0.129576	0.055090	2.352093	0.0187
SP500	-0.202205	0.233678	-0.865316	0.3869
ExcRate	-0.102141	0.600883	-0.169984	0.8650
VIX	-0.265981	0.071757	-3.706679	0.0002
@MONTH=2	0.275103	0.037808	7.276273	0.0000
@MONTH=3	0.090803	0.034368	2.642089	0.0082
@MONTH=4	0.130371	0.045525	2.863747	0.0042
@MONTH=5	0.317767	0.035883	8.855721	0.0000
@MONTH=6	-0.038888	0.040094	-0.969929	0.3321
@MONTH=7	0.150604	0.052972	2.843072	0.0045
@MONTH=8	0.497318	0.049309	10.08565	0.0000
@MONTH=9	0.400985	0.040483	9.905142	0.0000
@MONTH=10	0.080679	0.043075	1.872984	0.0611
@MONTH=11	0.048583	0.041991	1.156985	0.2473
@MONTH=12	0.036349	0.033668	1.079629	0.2803
LOG(SIGMA)	-3.426577	0.162590	-21.07491	0.0000
Regime 2				
C	0.027346	0.013450	2.033251	0.0420
BaaAaa	-0.130741	0.058517	-2.234251	0.0255
IndPro	0.271631	0.459410	0.591260	0.5543
IntRate	0.041284	0.064159	0.643457	0.5199
SP500	-0.306427	0.110342	-2.777064	0.0055
ExcRate	-1.686283	0.378430	-4.455996	0.0000
VIX	-0.062959	0.031434	-2.002874	0.0452
@MONTH=2	-0.039740	0.018112	-2.194094	0.0282
@MONTH=3	0.017400	0.021504	0.809143	0.4184
@MONTH=4	0.014413	0.020426	0.705641	0.4804
@MONTH=5	-0.024489	0.019840	-1.234341	0.2171
@MONTH=6	-0.034305	0.023764	-1.443552	0.1489
@MONTH=7	-0.004051	0.024557	-0.164952	0.8690
@MONTH=8	-0.030021	0.024582	-1.221271	0.2220
@MONTH=9	-0.029691	0.019330	-1.536009	0.1245
@MONTH=10	-0.010853	0.019424	-0.558749	0.5763
@MONTH=11	-0.035972	0.019804	-1.816367	0.0693
@MONTH=12	-0.019728	0.018595	-1.060914	0.2887
LOG(SIGMA)	-2.751370	0.045789	-60.08812	0.0000

The table shows the results of the Markov regime switching model for all the seven variables (in first differences) for regime 1 and 2 with the seasonality effects denoted as @MONTH=2 to @MONTH=12. Sample peridod 1986M01-2017M06.

Table B.5: Markov Regime Switching Regression Model continued

AR terms	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.192705	0.067089	2.872390	0.0041
AR(2)	-0.034387	0.065018	-0.528894	0.5969
AR(3)	0.049849	0.069946	0.712682	0.4760
AR(4)	-0.069524	0.065935	-1.054434	0.2917
Transition Prob	Coefficient	Std. Error	z-Statistic	Prob.
P11-C	0.275456	0.418566	0.658094	0.5105
P21-C	-2.904276	0.340473	-8.530130	0.0000
Mean dep var	0.003231	SD dep var		0.091189
SE of regression	0.085567	Sum squared resid		2.423486
DW stat	2.022347	Log likelihood		453.1887
AIC	-2.194041	SIC		-1.731442
HQC	-2.010349			

Continuation of the Markov regime switching model in table B.5, here with the AR terms. The corresponding transition probabilities matrix is here denoted as P11-C and P21-C. Sample period: 1986M01-2017M06.

Table B.6: Chow Breakpoint Test

Null Hypothesis: No breaks at specified breakpoints Varying regressors: All equation variables	
F-statistic Prob.F(18,341)	2.451141 0.0010
Log likelihood ratio Prob. Chi-Square(18)	45.87105 0.0003
Wald Statistic Prob. Chi-Square(18)	44.12053 0.0006

Table represents the Chow breakpoint test with the null hypothesis of no existing breakpoints. All variables are included (in first differences) in the test for the sample period of 1986M01-2017M06. The initial breakpoint is set to 2008M06.

C APPENDIX C - METHODS

C.1 Heteroscedasticity Tests

Breusch-Godfrey test

Breusch-Godfrey-Pagan tests serial correlation by deriving test statistics from the errors. The null hypothesis is no serial correlation.

$$
\begin{aligned}
H_0 &= p_0 + p_1 + p_2 + .. + p_k = 0 \\
H_1 &= p_0 \cup p_1 \cup p_2 \cup ... \cup p_k \neq 0
\end{aligned}
\tag{C.1}
$$

The test is initially for the dataset of a sample of residuals

$$
y_t = \beta_1 + \beta_2 x_2 t + ... + \beta_n x_n t + u_t \tag{C.2}
$$

Further, the residuals are regressed as shown in the following equation

$$
u_t = p_0 + p_1 y_{t-1} + ... + p_k u_{t-k} + \nu_t \tag{C.3}
$$

Breusch-Pagan test

Lastly, we consider the Breusch-Pagan test (BP). The null hypothesis, H_0 : assumes homoskedasticity

The test assumes the following regression

$$
y_t = \beta_0 + \beta_1 x_1 t + ... + \beta_n x_n t + u_t \tag{C.4}
$$

The residuals are further regressed on the independent variables, using a regression of the form

$$
\hat{u_t^2} = \delta_0 + \delta_1 \hat{Y}_t \tag{C.5}
$$

where $\hat{Y}_t$ is the predicted value from the original regression above and can be viewed as

$$\hat{Y}_t = \hat{\beta_0} + \beta_1\hat{x_1}t + ... + \hat{\beta_n}x_{nt} \tag{C.6}$$

The test evaluates the chi-squared statistic. Degrees of freedom are 1, and the test is based on R^2 and the number of observations. The chi-squared statistic is as follows

$$TR^2 \sim \chi^2(m) \tag{C.7}$$

$$\epsilon_t^2 = \beta_0 + \beta_1\epsilon_{t-1}^2 + \beta_2\epsilon_t^2 - 2 + ... + \beta_q\epsilon_t^2 - q + \nu_t \tag{C.8}$$

In this study we denote these two tests as the Breusch-Pagan-Godfrey test.

White test

The White test can be defined as follows. We estimate the following regression

$$y_t = b_1 + b_2x_t + b_3z_t + \epsilon_t \tag{C.9}$$

We can then describe the test statistic as this regression

$$\epsilon_t^2 = \alpha_0 + \alpha_1x_t + \alpha_3x_t^2 + \alpha_4z_t^2 + \alpha_5x_tz_t + \nu_t \tag{C.10}$$

The null hypothesis assumes homoskedasticity, and if the χ^2 is greater than the critical χ^2 value, then we reject the null hypothesis and conclude that the series are heteroskedastic.

HAC (Newey-White)

In our study we use the extension of the Newey-White method to correct for heteroscedasticity. This can be shown as

$$\hat{\sum}_{NW} = (XX)^{-1} T \hat{\Lambda} (XX)^{-1} \tag{C.11}$$

where the LRCOV estimator is shown as $\hat{\Lambda}$, denoted as the Long-run-Covariance estimation. The latter will not be further described here.

Residual Sum of Squares

RSS denotes the residual sum of squares and gives an estimation between the model and the data. It can be described as

$$RSS = \sum_{t=1}^{T} e_t^2 = \sum_{t=1}^{T} (y_t - \hat{y}_t)^2 \tag{C.12}$$

In this equation, e is the given residual and Y and $\hat{Y}$ is the observation and the predicted value.

C.2 Methods for Model Evaluation In-Sample

To assess the two models used in this study (VAR and MS), we perform forecast evaluation. Specifically, the evaluation metrics considered are the root-mean-squared error (RMSE), the mean absolute error(MAE), the mean absolute percentage error (MAPE) and the mean absolute scaled error (MASE). There are drawbacks to all these metrics, but evaluating several of them can help overcome their limitations and give us an overall understanding of the performance of the applied methods. The models can be formulated as follows

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2} \tag{C.13}$$

$$\mathrm{MAE} = \frac{\sum_{t=1}^{T}|y_t - \hat{y}_t|}{T} \tag{C.14}$$

$$\mathrm{MAPE} = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{y_t - \hat{y_t}}{y_t}\right| \tag{C.15}$$

$$\mathrm{MASE} = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{|y_t - \hat{y_t}|}{\frac{1}{T-m}\sum_{t=m+1}^{T}|y_t - \hat{n}_{t-m}|}\right) \tag{C.16}$$

Random Walk

In addition to the evaluation metrics, the two models are also compared with two benchmark models. The first one being the random walk (RW). This model can be defined as

$$y_t = t_{t-1} + a + \epsilon_t \tag{C.17}$$

The white noise term is ϵ_t, with a zero mean, and a variance equal to one

ARIMA

In addition to the Random Walk, we also include the Autoregressive Integrated Moving Average (ARIMA) model. The forecast of ARIMA consists of stationarized series, its lags, the lags of the forecast errors and a constant. The ARIMA model can be described as an ARIMA(p, d, q) model by the following equation

$$(1 - \sum_{k=1}^{P} \alpha_k L^k)(1 - d)\chi_t = (1 + \sum_{k}^{q} = 1\beta_k L^k)\epsilon_t \tag{C.18}$$

From this equation the L is the lag operator, X_t is the variable, and $\epsilon_1, ..., \epsilon_k$ are the errors. $\alpha_1, ..., \alpha_p$ and $\beta_1, ..., \beta_p$ are the autoregressive and moving average terms. The rest of equation can be explained as follows

- d is the number of differences needed to obtain stationarity
- q is the number of lagged forecasts errors in the predicted equation
- p is the number of autoregressive terms

D APPENDIX D - MODEL EVALUATION

D.1 Model Evaluation In-Sample

In this section the results obtained from the forecast evaluation and the benchmark models are analyzed and evaluated in order to investigate the performance of VAR and RS model. The results are reported in table D.1. The evaluation was done in-sample. As explained by Hyndman et al. (2007), there should be applied two models for prediction, a naive, Random Walk, and ARIMA, here, specified as ARIMA(p, d, q). Here, the evaluation metrics such as RMSE, MAE, MAPE, MASE are considered and denoted in percentage (%) in table D.1. Correlation is here denoted as CORR and is between the predicted values of all the model against the actual values of oil price (in first differences).

We prefer the RMSE metric to be as small as possible, as it indicates a better fit of a model. As shown in table D.1, the RMSE of VAR is close to 0.034 and for RS around 0.046. As both models show low values of the RMSE, they perform well. However, VAR shows a better fit according to the results. Lower values of MAE is also preferred. The MAE of MS and VAR is 0.034 and 0.026, respectively. VAR is lower and thus, forecasts better than MS. As seen in table D.1, MASE for MS is around 0.40 and around 0.31 for VAR. We prefer to have a MASE below 1 as it indicates that the forecast is better than a naive forecast. Both models evaluations have a MASE below 1, with the best performance of VAR, according to this metric.

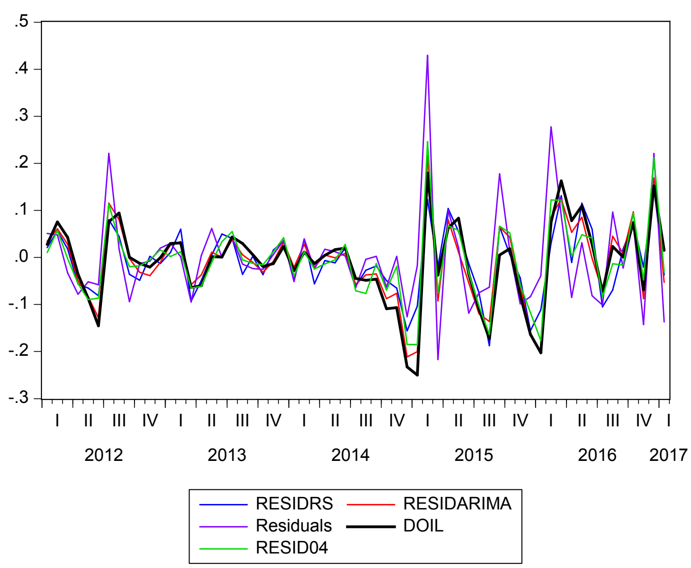
The prediction accuracy is also considered by evaluating the MAPE. A drawback of MAPE is the need to take into consideration if there are zero values present in the time series, which would imply a division by zero. Data observations close to zero, can also result in extreme MAPE values. As observed in the logarithmic forms of the oil price, there is one zero value. This is adjusted by calculating the mean of the oil price value the day before and after to replace the zero value. As this was done after performing both the VAR and MS model, this has no effect on our main results. The RS model has a MAPE of around 203%, whereas VAR is around 166%.

To evaluate the performance, we also apply the RW and ARIMA model. The forecast metrics for the RW naive model are the highest for all the models evaluated. RMSE is twice as high as for RS and VAR. The same is evident for the MAE value. The MAPE is shown to be 438%, which is twice as high as for RS and almost three times the VAR and ARIMA. The correlation metric is also the lowest for the RW, with a value of 0.609169, indicating a less good fit with the actual values of oil series (in first differences) compared to the VAR and RS estimation. As, it can be seen the second benchmark model seems to perform slightly better than the VAR model. The ARIMA model has the lowest values compared to all the models and the highest CORR value, which indicates that this model performs best. As seen from the table D.1, the results are only slightly better than the performance of the VAR model, therefore we conclude that VAR is good estimator.

Table D.1: Model Evaluation Metrics

	RMSE	MAE	MAPE	MASE	CORR
MS	0.045852	0.033519	202.5582	0.399146	0.864793
VAR	0.033587	0.025999	166.1281	0.305778	0.931345
RW	0.092637	0.069630	438.9969	0.820116	0.609169
ARIMA	0.030198	0.022811	115.3799	0.26953	0.945564

Figure D.1: Forecast Evaluation of MS, ARIMA, VAR and RW to Oil



Forecast evaluation for the residuals of all the four models against the actual value of oil price (in first differences) denoted as the black DOIL line. RESIDRS (residuals of the MS model), RESIDARIMA (residuals of the ARIMA), RESID04 is the best fitted residual of the VAR model and Residuals (residuals from the RW model). The period sample here shows only 2006M01-2017M01 for illustrative purposes.

Figure D.2: RW Forecast Evaluation to Oil

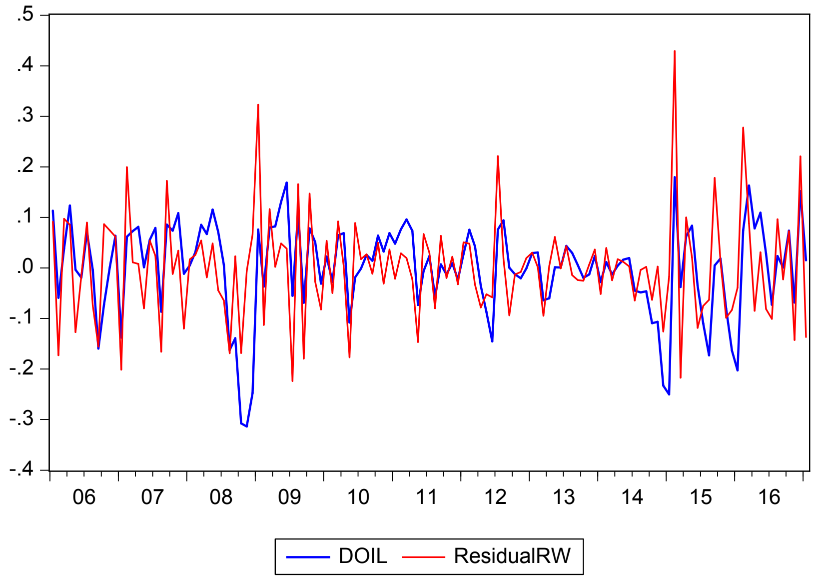
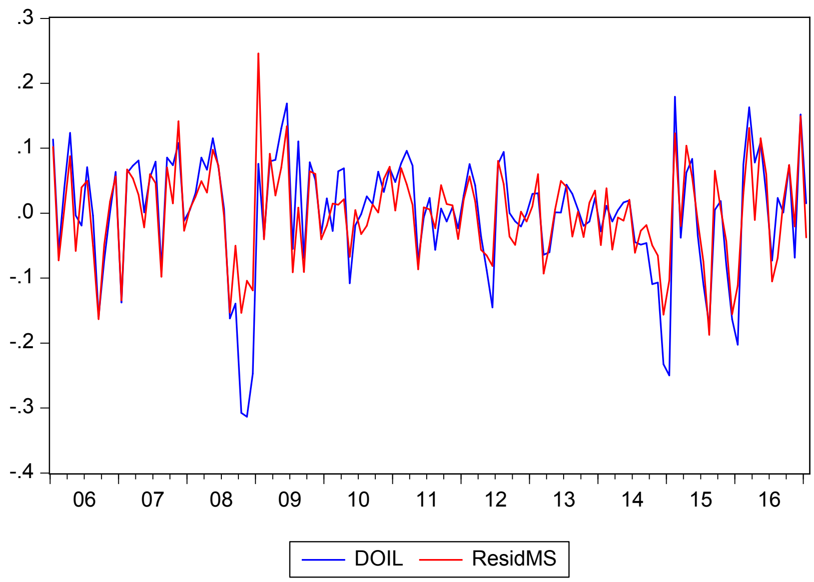


Figure D.3: MS Forecast Evalutaion to Oil



Forecast evaluation in table D.2 for the residuals of the MS model (ResidMS) against the actual values of the oil price (in first differences), and RW model residuals (ResidualRW) against the actual value of oil price (in first differences) in table D.3. The period sample is 2006M01-2017M01. Only thw period sample 2006M01-2017M01 is displayed here for illustrative purposes.

Figure D.4: VAR Forecast Evaluation

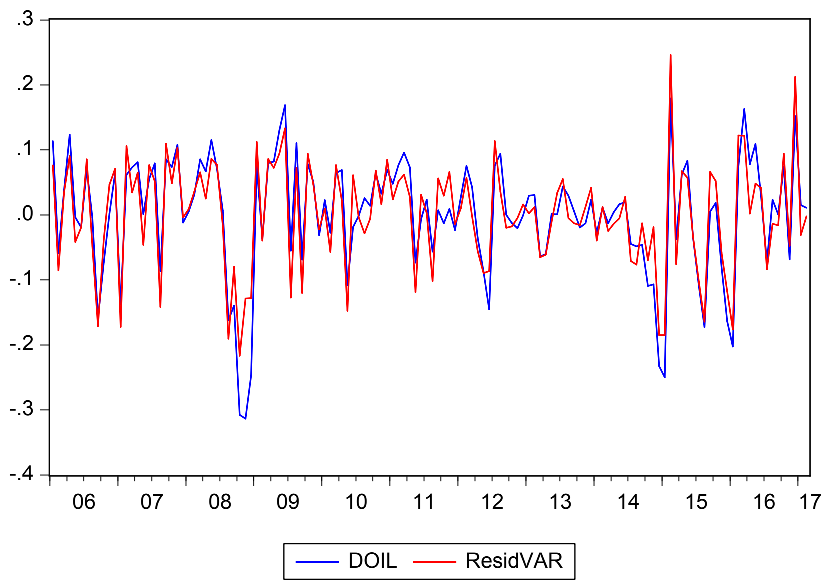


Figure D.5: ARIMA Forecast Evaluation to Oil

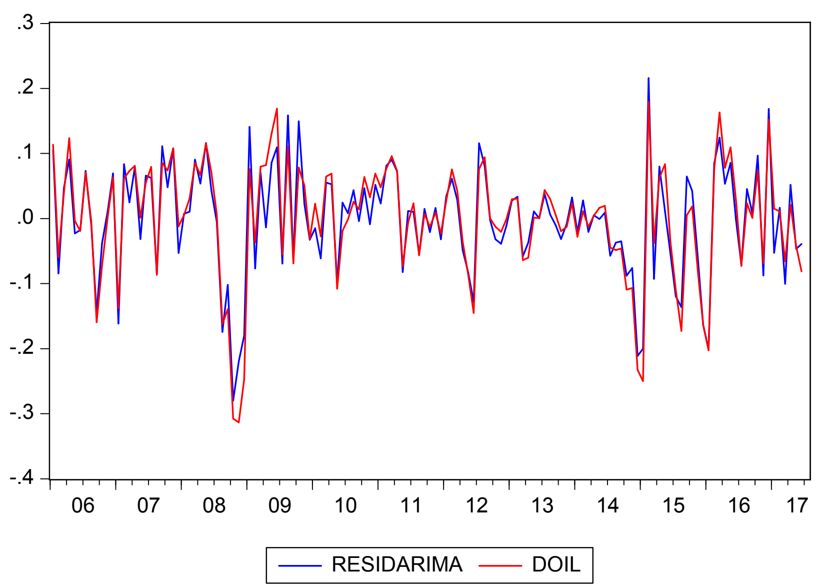


Table D.4 displays forecast evaluation for the residuals of the VAR (ResidVAR) as the values of RESID04, best fitted residuals of the VAR model. Table D.5 displays the residuals of the ARIMA model against the actual values of oil price (in first differences). The period sample is 2006M01-2017M01. The period sample 2006M01-2017M01 is displayed here for illustrative purposes.

Table D.2: Forecast Evaluation: VAR Model

Forecast	F-stat	F-prob	
RESID01	106.4169	0.0000	
RESID02	374.8072	0.0000	
RESID03	93.81565	0.0000	
RESID04	0.000000	1.0000	
RESID05	205.7459	0.0000	
RESID06	385.5438	0.0000	
RESID07	81.42365	0.0000	
Evaluation statistics			
Forecast	RMSE	MAE	height
	RESID01	0.120980	0.090925
RESID02	0.091766	0.069585	
RESID03	0.115955	0.085586	
RESID04	0.033537	0.025946	
RESID05	0.100324	0.076142	
RESID06	0.094534	0.071949	
RESID07	0.181068	0.133539	

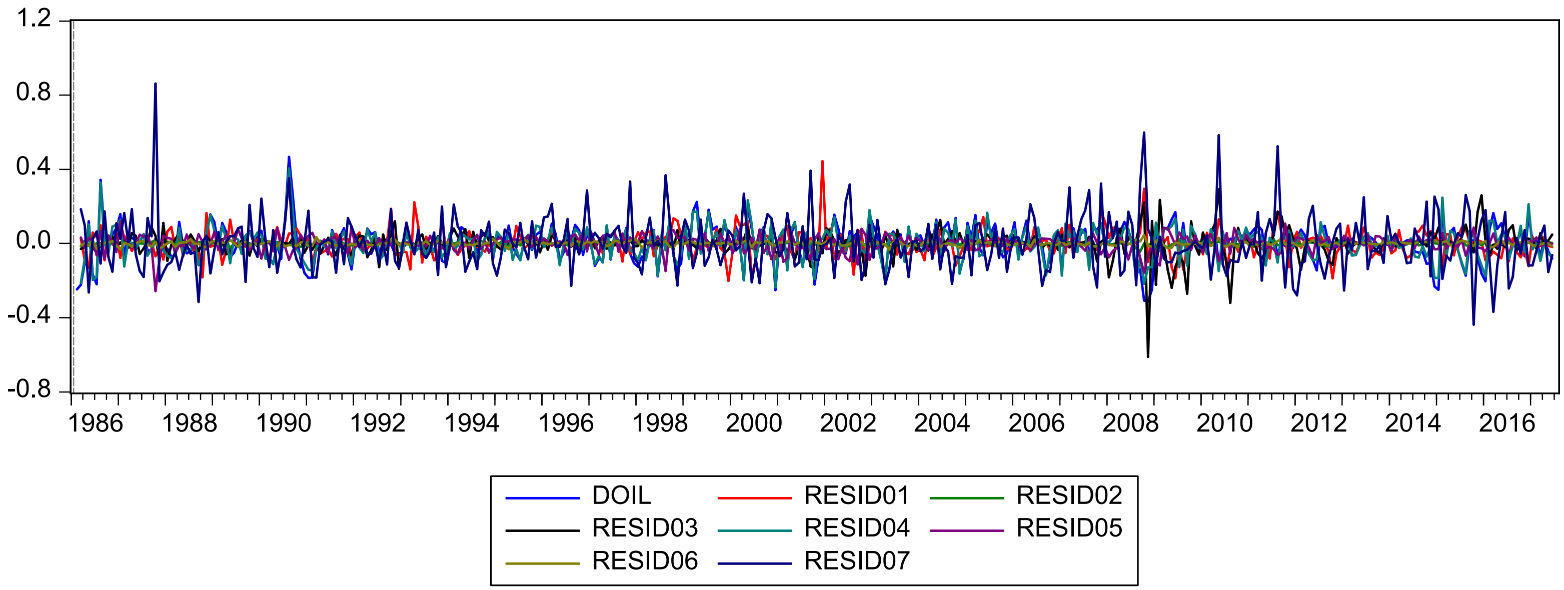
Statistical forecast evaluation of the VAR model, represented by RMSE and MAE. RESID01 - RESID07 denoted as the corresponding residuals of all the seven variables included in the model. Sample period: 1986M01-2017M06

Table D.3: Forecast Evaluation: MS Model

Evaluation Statistics Forecast	RMSE	MAE
RESIDMS	0.045852	0.033519

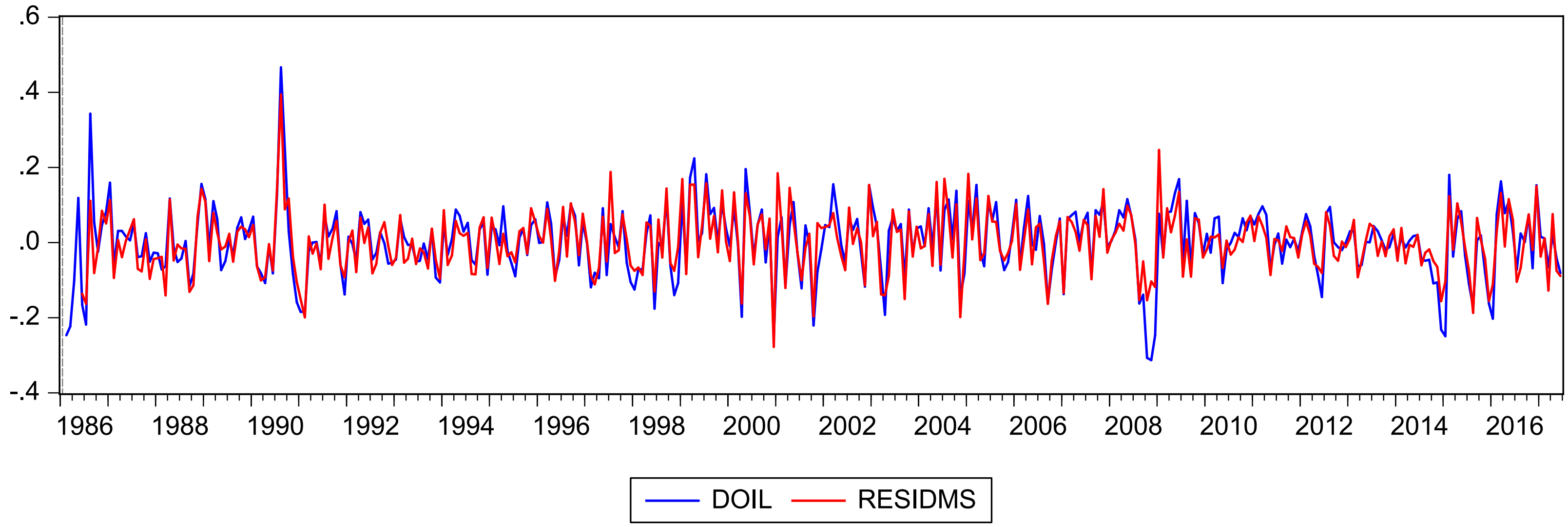
Forecast evaluations for the Markov regime switching modeThe RMSE and MAE statistics are represented. RESIDMS denotes the residuals for the RS model. Sample period: 2006M01-2017M01

Figure D.6: VAR Forecast



VAR forecast for the residuals of the VAR model compared to the first difference of Oil. The period sample is 1986M01-2017M01

Figure D.7: MS Forecast



MS forecast for the residual of the MS model against the first difference of Oil. The period sample is 1986M01-2017M01

REFERENCES

Aastveit, Knut Are, Hilde C Bjørnland, and Leif Anders Thorsrud. 2015. What drives oil prices? emerging versus developed economies. *Journal of Applied Econometrics* 30(7):1013–1028.

Akram, Q Farooq. 2009. Commodity prices, interest rates and the dollar. *Energy economics* 31(6):838–851.

Alexander, Carol. 2008. *Market risk analysis, practical financial econometrics.* Wiley.

Alquist, Ron, and Olivier Gervais. 2013. The role of financial speculation in driving the price of crude oil. *The Energy Journal* 35–54.

Balcilar, Mehmet, Rangan Gupta, and Stephen M Miller. 2015. Regime switching model of us crude oil and stock market prices: 1859 to 2013. *Energy Economics* 49: 317–327.

Barsky, Robert B, and Lutz Kilian. 2004. Oil and the macroeconomy since the 1970s. *Journal of Economic Perspectives* 18(4):115–134.

Baumeister, Christiane, and Lutz Kilian. 2016. Understanding the decline in the price of oil since june 2014. *Journal of the Association of Environmental and Resource Economists* 3(1):131–158.

Bernanke, Ben. 2016. The relationship between stocks and oil prices. *Ben Bernanke's Blog on Brookings posted on February* 19.

Bjørnland, Hilde, Vegard Larsen, and Junior Maih. 2017. Oil and macroeconomic (in) stability.

Bjørnland, Hilde C. 2000a. The dynamic effects of aggregate demand, supply and oil price shocks- a comparative study. *The Manchester School* 68(5):578–607.

Bjørnland, Hilde C. 2000b. Var models in macroeconomic research. *Statistisk Sentralbyrå* 1–29.

Bjørnland, Hilde Christiane, and Leif A. Thorsrud. 2015. *Applied time series for macroeconomics.* Gyldendal.

Blanchard, Olivier J, and Marianna Riggi. 2013. Why are the 2000s so different from the 1970s? a structural interpretation of changes in the macroeconomic effects of oil prices. *Journal of the European Economic Association* 11(5):1032–1052.

Broadstock, David C, and George Filis. 2014. Oil price shocks and stock market returns: New evidence from the united states and china. *Journal of International Financial Markets, Institutions and Money* 33:417–433.

Brooks, Chris. 2008. *Introductory econometrics for finance 2nd edition*. Cambridge University Press.

Burbidge, John, and Alan Harrison. 1984. Testing for the effects of oil-price rises using vector autoregressions. *International Economic Review* 459–484.

Choi, Kyongwook, and Shawkat Hammoudeh. 2010. Volatility behavior of oil, industrial commodity and stock markets in a regime-switching environment. *Energy Policy* 38(8):4388–4399.

Dickey, David A., and Wayne A. Fuller. 1972. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74:427–431.

Diebold, FX. 2015. Forecasting in economics, business, finance and beyond. *Department of Economics, University of Pennsylvania*.

ECB. 2014. Marcel Fratzscher, Daniel Schneider and Ine Van Robays. Oil Prices, Exchange Rates and Asset Prices. `https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1689.pdf?06716196e143159bdcc5d65b2461f68d`. [Accessed 2018-05-26].

Ehrmann, Michael, Martin Ellison, and Natacha Valla. 2003. Regime-dependent impulse response functions in a markov-switching vector autoregression model. *Economics Letters* 78(3):295–299.

EIA. 2017. International energy outlook 2017. `https://www.eia.gov/outlooks/ieo/`. [Accessed: 2018-05-31].

EIA. 2018. Oil: Crude and petroleum products explained oil prices and outlook. `https://www.eia.gov/energyexplained/index.php?page=oil_prices`. [Accessed: 2018-06-04].

Eichenbaum, Martin, and Charles L Evans. 1995. Some empirical evidence on the effects of shocks to monetary policy on exchange rates. *The Quarterly Journal of Economics* 110(4):975–1009.

Elliott, Graham, Thomas J. Rothenberg, and James H. Stock. 1992. Efficient tests for an autoregressive unit root. *Technical Working Paper* 130:1–36.

Fong, Wai Mun, and Kim Hock See. 2002. A markov switching model of the conditional volatility of crude oil futures prices. *Energy Economics* 24(1):71–95.

Fratzscher, Marcel, Daniel Schneider, and Ine Van Robays. 2014. Oil prices, exchange rates and asset prices.

Gertler, Mark, and Cara S Lown. 1999. The information in the high-yield bond spread for the business cycle: evidence and some implications. *Oxford Review of Economic Policy* 15(3):132–150.

Granger, Clive WJ. 1996. Can we improve the perceived quality of economic forecasts? *Journal of Applied Econometrics* 455–473.

Hamilton, James D. 1983. Oil and the macroeconomy since world war ii. *Journal of political economy* 91(2):228–248.

Hamilton, James D. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society* 357–384.

Hamilton, James D. 1996. This is what happened to the oil price-macroeconomy relationship. *Journal of Monetary Economics* 38(2):215–220.

Hamilton, James D. 2003. What is an oil shock? *Journal of econometrics* 113(2): 363–398.

Hamilton, James D. 2005. Oil and the macroeconomy. *The new Palgrave dictionary of economics* 2.

Hamilton, James D. 2008. Understanding crude oil prices. *JSTOR* 30:179–206.

Hamilton, James D, and Dong Heon Kim. 2000. A re-examination of the predictability of economic activity using the yield spread. Tech. Rep., National Bureau of Economic Research.

Hammoudeh, Shawkat, Sel Dibooglu, and Eisa Aleisa. 2004. Relationships among us oil prices and oil industry equity indices. *International Review of Economics & Finance* 13(4):427–453.

Holm-Hadulla, Fédéric, and Kirstin Hubrich. 2017. Macroeconomic implications of oil price fluctuations: a regime-switching framework for the euro area.

Hyndman, Rob J, Andrey V Kostenko, et al. 2007. Minimum sample size requirements for seasonal forecasting models. *Foresight* 6(Spring):12–15.

Johansen, Søren. 1991. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica: Journal of the Econometric Society* 1551–1580.

Kaufmann, Robert K. 2011. The role of market fundamentals and speculation in recent price changes for crude oil. *Energy Policy* 39(1):105–115.

Kilian, Lutz, and Bruce Hicks. 2013. Did unexpectedly strong economic growth cause the oil price shock of 2003–2008? *Journal of Forecasting* 32(5):385–394.

Kilian, Lutz, and Thomas K Lee. 2014. Quantifying the speculative component in the real price of oil: The role of global oil inventories. *Journal of International Money and Finance* 42:71–87.

Kilian, Lutz, and Daniel P Murphy. 2014. The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics* 29(3): 454–478.

Kilian, Lutz, and Cheolbeom Park. 2009. The impact of oil price shocks on the us stock market. *International Economic Review* 50(4):1267–1287.

Krolzig, Hans-Martin. 1997. *Markov-switching vector autoregressions: Modelling, statistical inference, and application to business cycle analysis*, vol. 454. Springer Science & Business Media.

Kwiatkowski, Denis, Peter C.B. Phillips, Peter Schmidt, and Yongcheol Shin. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics* 54: 159–178.

Liew, Venus Khim-Sen. 2004. Which lag length selection criteria should we employ?

Lippi, Francesco, and Andrea Nobili. 2012. Oil and the macroeconomy: a quantitative structural analysis. *Journal of the European Economic Association* 10(5):1059–1083.

Lizardo, Radhamés A, and André V Mollick. 2010. Oil price fluctuations and us dollar exchange rates. *Energy Economics* 32(2):399–408.

Mork, Knut Anton. 1989. Oil and the macroeconomy when prices go up and down: an extension of hamilton's results. *Journal of political Economy* 97(3):740–744.

Newey, Whitney K, and Kenneth D West. 1987. Hypothesis testing with efficient method of moments estimation. *International Economic Review* 777–787.

Nkomo, JC. 2017. The impact of higher oil prices on southern african countries. *Journal of Energy in Southern Africa* 17(1):10–17.

Nomikos, Nikos K, and Panos K Pouliasis. 2011. Forecasting petroleum futures markets volatility: The role of regimes and market conditions. *Energy Economics* 33(2):321–337.

Oxford Energy. 2001. Robert Mabro, Oxford Energy. `https://www.oxfordenergy.org/publications/does-oil-price-volatility-matter/`. [Accessed 04-June-2018].

Park, Jungwook, and Ronald A Ratti. 2008. Oil price shocks and stock markets in the us and 13 european countries. *Energy economics* 30(5):2587–2608.

Perron, Pierre, et al. 2006. Dealing with structural breaks. *Palgrave handbook of econometrics* 1(2):278–352.

Phillips, Peter C. B., and Pierre Perron. 1988. Testing for a unit root in time series regression. *Biometrika* 75:335–346.

Pindyck, Robert S, and Julio J Rotemberg. 1988. The excess co-movement of commodity prices.

Ryan, Laura, and Bronwen Whiting. 2017. Multi-model forecasts of the west texas intermediate crude oil spot price. *Journal of Forecasting* 36(4):395–406.

Sadorsky, Perry. 1999. Oil price shocks and stock market activity. *Energy economics* 21(5):449–469.

Salisu, Afees A, and Ismail O Fasanya. 2013. Modelling oil price volatility with structural breaks. *Energy Policy* 52:554–562.

Sims, Christopher A. 1980. Macroeconomics and reality. *Econometrica: Journal of the Econometric Society* 1–48.

Tang, Weiqi, Libo Wu, and ZhongXiang Zhang. 2010. Oil price shocks and their short-and long-term effects on the chinese economy. *Energy Economics* 32:S3–S14.

The New York Times. 2017. Oil prices: What to make of the volatility, krauss, clifford. `https://www.nytimes.com/interactive/2017/business/energy-environment/oil-prices.html`. [Accessed: 2018-06-02].

The World Bank. 2013. Jun Erik Rentschler. Oil Price volatility - its risk on economic growth and development. `http://blogs.worldbank.org/developmenttalk/oil-price-volatility-its-risk-economic-growth-and-development`. [Accessed 04-June-2018].

Vo, Minh T. 2009. Regime-switching stochastic volatility: Evidence from the crude oil market. *Energy Economics* 31(5):779–788.

Zhu, Huiming, Xianfang Su, Wanhai You, and Yinghua Ren. 2017. Asymmetric effects of oil price shocks on stock returns: evidence from a two-stage markov regime-switching approach. *Applied Economics* 49(25):2491–2507.